(19) Europäisches Patentamt
Eur pean Patent Offic
Office européen des brevets

(11) Publication number: **0 528 906 B1**

(12) # EUROPEAN PATENT SPECIFICATION

(45) Date of publication of patent specification: **11.10.95**

(51) Int. Cl.[6]: **C12N 15/53**, C12N 9/02, C12N 1/21, C12N 5/10, C12Q 1/32

(21) Application number: **91909339.3**

(22) Date of filing: **25.04.91**

(86) International application number:
PCT/US91/02866

(87) International publication number:
WO 91/17251 (14.11.91 91/26)

(54) **STEROID 5-g(a)-REDUCTASES.**

(30) Priority: **30.04.90 US 517661**

(43) Date of publication of application:
03.03.93 Bulletin 93/09

(45) Publication of the grant of the patent:
11.10.95 Bulletin 95/41

(84) Designated Contracting States:
AT BE CH DE DK ES FR GB GR IT LI LU NL SE

(56) References cited:

PROCEEDINGS OF THE NATL. ACADEMY OF SCIENCES USA, vol. 87, May 1990; S. ANDERSSAN et al., pp. 3640-3644

CHEMICAL ABSTRACTS, vol. 103, 05 August 1985, Columbus, OH (US); B. HOUSTON et al., p. 316, no. 35248j

CHEMICAL ABSTRACTS, vol. 112, 04 June 1990, Columbus, OH (US); S. ANDERSSON et al., p. 145, no. 211718j

CHEMICAL ABSTRACTS, vol. 99, 01 August 1983, Columbus, OH (US); KADOHAMA et al., p. 68, no. 33243d

(73) Proprietor: **BOARD OF REGENTS THE UNIVERSITY OF TEXAS SYSTEM**
201 West 7th Street
Austin, Texas 78701 (US)

(72) Inventor: **ANDERSSON, Stefan**
124 W. 60th Street
19M New York, NY 10023 (US)
Inventor: **RUSSELL, David, W.**
1541 W. Colorado
Dallas, TX 75208 (US)

(74) Representative: **Grünecker, Kinkeldey, Stockmair & Schwanhäusser Anwaltssozietät**
Maximilianstrasse 58
D-80538 München (DE)

EP 0 528 906 B1

Description

The government may own certain rights in the present invention pursuant to NIH grant GM 43753 and HL 20948.

The present invention relates generally to enzymes, termed steroid 5α-reductases, which function biologically to catalyze the conversion of testosterone to dihydroxytestosterone. In particular aspects, the invention relates to the preparation of this enzyme from various sources by recombinant techniques, to nucleic acid segments which encode the enzyme or which can be used as probes for the selection of related sequences, as well as to assay methods for the identification of candidate substances which will affect the activity of the enzyme.

The enzyme steroid 5α-reductase is a microsomal protein that plays a central role in human sexual differentiation and androgen physiology. Interest in this protein arises from several distinguishing characteristics. First, steroid 5α-reductase catalyzes the conversion of testosterone into the more potent androgen dihydrotestosterone (1). This latter steroid induces a program of differentiation during fetal development that leads to the development of the male external genitalia (2). Second, mutations in the gene for steroid 5α-reductase give rise to a rare form of male pseudohermaphroditism in which affected males develop normal internal urogenital tracts but fail to develop external male structures (3). Third, the expression of the gene is regulated by androgens in tissues such as the prostate and liver (4). A fourth distinguishing feature of steroid 5α-reductase is its role in several endocrine abnormalities including benign prostate hyperplasia, male pattern baldness, acne, and hirsutism (5-7).

It is this fourth role which has led researchers towards the development of agents that will serve to inhibit the enzyme, with the hope that such agents will prove useful in the treatment of one or more of these conditions (8,9). Since the product of steroid 5α-reductase activity, dihydrotestosterone, is involved in inducing these and perhaps other conditions, it is believed that by inhibiting steroid 5α-reductase action, one can ameliorate one or more aspects of the particular condition. The drugs which have been used a therapeutic agents include principally 4-azasteroid derivatives such as MK-906 (Finasteride) and 4-MA (8,9) that function as competitive inhibitors of the enzyme (10). The exact mechanism by which these compounds act in vivo has yet to be elucidated.

While these competitive inhibitors of steroid 5α-reductase have shown some promise, e.g., in the treatment of benign prostatic hyperplasia, in general, these agents appear to suffer from a variety of problems and potential drawbacks, including limited efficacy and even hepatotoxicity. Furthermore, the development of additional inhibitors has been greatly hampered due to the previous lack of a useful, relatively simple test system which can be used to screen for new inhibitors. The previous lack of knowledge in the art concerning the enzyme itself, such as knowledge about its structure, has hampered the development of new therapeutic agents. Efforts in this regard have been hampered by the very low levels of expression of this enzyme in most tissues, even in tissues which are responsive to androgens (10-12).

Accordingly, if medical science is to succeed in the development of novel and more efficacious steroid 5α-reductase inhibitors, there is currently a great need for an expansion of our knowledge of this enzyme. There is a great need to develop improved methods for screening for compounds which affect the function of the enzyme in one or more ways, such as highly sensitive and rapid screening methods which can be applied to screen for such agents from a large panel of candidate substances. Moreover, there is a great need to provide means for preparing improved compositions of biologically active steroid 5α-reductase, particularly human steroid 5α-reductase, which can be employed in the furthering of our understanding of the enzyme as well as in the development of screening protocols.

The present invention concerns, in a general sense, compositions and methods for the synthetic preparation of human steroid 5α-reductases, as well as their biological functional equivalents, and to methods employing these species in the identification of candidate substances capable of inhibiting or otherwise modifying their enzymatic function.

In certain general and overall embodiments, therefore, the invention concerns recombinant vectors and isolated DNA segments encoding a steroid 5α-reductase. DNA segments of the invention may also encode biologically functional equivalent proteins or peptides which have variant amino acids sequences, such as species which incorporate changes based on considerations such as the relative hydropathic score of the amino acids being exchanged.

As used herein, the term "DNA segment" in intended to refer to a DNA molecule which has been isolated free of total genomic DNA of a particular species. Therefore, a DNA segment encoding a steroid 5α-reductase is intended to refer to a DNA segment which contains such coding sequences yet is isolated away from total genomic DNA of the species from which the DNA is obtained. Included within the term

"DNA segment", are DNA segments which may be employed in the preparation of vectors, as well as the vectors themselves, including, for example, plasmids, cosmids, phage, viruses, etc.

In the context of the present invention, the term "steroid 5α-reductase" is intended to refer to any protein or peptide having the biological or immunological identity, or both, of a human steroid 5α-reductase enzyme as exemplified, e.g., by functional equivalents.

In more particular embodiments, the invention concerns recombinant vectors and isolated DNA segments incorporating DNA sequences which encode a steroid 5α-reductase that incudes within its amino acid sequence the amino acid sequence of Figure 7, corresponding to human steroid 5α-reductase. Recombinant vectors and isolated segments may therefore variously include the human steroid 5α-reductase coding region itself, coding regions bearing selected alterations or modifications in the basic coding region or may encode larger proteins which nevertheless include sequences which will confer steroid 5α-reductase activity. Furthermore, and in any event, it should be appreciated that due to codon redundancy and functional equivalency this aspect of the invention is not limited to the particular DNA sequences shown in Figure 7.

Recombinant vectors such as the foregoing are useful both as a means for preparing quantities of the enzyme, and as a means for preparing shorter peptides. It is contemplated that where steroid 5α-reductase proteins of the invention are made by recombinant means, one may employ either prokaryotic or eukaryotic expression systems.

Where expression of a steroid 5α-reductase enzyme in a host is contemplated, it may be desirable to employ a vector, such as a plasmid, that incorporates an origin of replication, as exemplified by the eukaryotic vectors of the pCMV series, like pCMV4. Additionally, for the purposes of expression in host systems, one will desire to position the coding sequences adjacent to and under the control of an effective eukaryotic promoter, such as an SV40 or CMV promoter in eukaryotic systems. To bring a coding sequence under the control of such a promoter, whether it be a eukaryotic or prokaryotic promoter, all that is generally needed is to position the 5' end of the transcription initiation site of the transcriptional reading frame of the protein between about 1 and about 50 nucleotides "downstream" of (i.e., 3' of) with respect to the promoter chosen.

Furthermore, where host expression is contemplated, one will typically desire to incorporate into the transcriptional unit which includes the enzyme, an appropriate polyadenylation site (e.g., 5'-AATAAA-3') i eukaryotes, or a transcriptional terminator in the case of prokaryotes. Typically, the poly A addition site is placed about 30 to 2000 nucleotides "downstream" of the termination site of the protein at a position prior to transcription termination. A similar positioning of the prokaryotic terminator is also typical.

Useful eukaryotic vectors which include all of the foregoing, and into which the gene of the present invention can be inserted with little difficulty are well known. For example, suitable eukaryotic vectors include pCD and pCMV, with the most preferred system being pCMV. In addition to pCD and pCMV vectors, other preferred eukaryotic expression vectors include pMSG and pSVL from Pharmacia LKB Technology, Piscataway, N.J. These utilize the MMTV and SV40 late promoters, respectively. A DNA, such as shown in Fig. 7, can readily be inserted into one of the foregoing vectors via the Eco RI restriction site "upstream" of (i.e. 5' of) the initiation codon (ATG) that begins translation of the encoded enzyme.

It is contemplated that virtually any of the commonly employed eukaryotic host cells can be used in connection with steroid 5α-reductase expression in accordance herewith. Examples include cell lines typically employed for eukaryotic expression such as AtT20, HepG2, VERO, HeLa, CHO, WI 38, BHK, COS-7, RIN and MDCK cell lines. A preferred line for use in eukaryotic expression embodiments of the present invention is the COS-7 system. Of course, where eukaryotic hosts are employed, it is known that recombinant sequences may be either maintained extrachromosomally, or may be actually incorporated or integrated into the genome of the host cell. For long term expression, it will generally be preferred to employ systems wherein genomic integration is achieved, such as CHO or HepG2. However, where mere transient expression is desired, such as for recombinant screening purposes, extrachromosomal transformation may be sufficient, such as exemplified by COS-7 or HeLa cells.

Prokaryotic expression is an alternative which can be employed where desired. Typically, prokaryotic promoters which may be employed include P_L, T7 and lac promoter, with T7 being generally preferred. Other preferred bacterial expression vectors include plasmid pKK233-2 and pKK233-3, available from Pharmacia LKB Technology. These utilize the tac and trc promoters, respectively.

Of course, even where a eukaryotic hook-up and expression is used, one will nevertheless usually desire to include a prokaryotic origin of expression, as well as selective markers operabl in prokaryotic systems, to allow "shuttling" of sequences from construction in prokaryotic to expression in eukaryotic.

In certain embodiments of the invention it is contemplated that DNA fragments both shorter and longer which incorporate sequences from Figure 7, or related sequences, will find additional utilities, including uses

in the preparation of short enzymatically active peptide or even as short DNA fragment hybridization probes for use, e.g., in screening clone banks. In any event, fragments corresponding to the Figure 7 sequence for stretches of as short as 10 or so nucleotides, will find utility in accordance with these or other embodiments. By having stretches of at least about 10 to 20 nucleotides in common with the disclosed DNA sequence of Figure 7, or its complement, a DNA segment will have the ability to form a preferential hybrid with steroid 5α-reductase DNA, particularly under more stringent conditions such as 0.15M NaC1 and 0.02M sodium citrate pH 7.4 at 50 °C. While a complementary or common stretch of about 10 or so nucleotides will ensure the ability to form a stable hybrid, longer stretches of complementarily may prove more desirable for certain uses. Thus, one may desire to certain uses DNA segments incorporating longer stretches of complementarily, for example, on the order of 18, 22 or even 25 or so bases.

An important aspect of the invention concerns a method for the production of steroid 5α-reductase by recombinant means, as well as use of the recombinantly produced enzyme in screening assays. Screening assays of the present invention will generally involve determining the ability of a candidate substance to affect the enzymatic activity of the enzyme, such as the screening of candidate substances to identify those that will inhibit or otherwise modify its enzymatic function. Typically, this method will include recombinantly preparing steroid 5α-reductase, followed by testing the recombinant steroid 5α-reductase with a candidate substance to determine the ability of the substance to affect its enzymatic function. Due to significant differences which have been identified by the inventors between the human enzyme and enzymes of other species, one will typically prefer to employ the human enzyme in connection with these screening methods where one intends to identify candidate substances for use in humans.

In a typical screening assay for identifying candidate substances, one may desire to employ the same recombinant expression host as the starting source for obtaining the enzyme, generally prepared in the form of a crude homogenate. Recombinant cells expressing the enzyme may be washed and homogenized to prepare a crude protein homogenate in a desirable buffer such as disclosed herein. In a typical assay, an amount of protein from the cell homogenate, such as 10 to 50 μg of cell homogenate protein, is placed into a small volume, e.g., 0.5 ml, of an appropriate assay buffer (e.g., 0.1 M potassium phosphate buffer, pH 6.6 rat enzyme, pH 7.0, human enzyme). Steroid substrates, such as testosterone, progesterone or androstenedione, are added to the admixture in convenient amounts, such as, e.g., 0.1 to 20 μM, and the reaction allowed to initiate by the addition of the cofactor NADPH. Where one uses an appropriate known substrate for the enzyme, one can, in the foregoing manner, obtain a baseline activity for the recombinantly produced enzyme. Then, to test for inhibitors or modifiers of the enzyme function, one can incorporate into the admixture a candidate substance whose effect on the enzyme is to be tested. By comparing reactions which are carried out in the presence or absence of the candidate substance, one can then obtain information regarding the effect of the candidate substance on the normal enzymatic function of the enzyme.

In preferred assays, the enzymatic function is measured by simply measuring the amount of product produced, or substrate used up, in the experimental reaction versus the control over a period of time. One may find it of benefit, therefore, to measure the rate at which a particular substrate is used, or product appears. In any event, the inventors have found that a convenient method for measuring the disappearance of substrate or appearance of product is through the use of a labeled substrate, such as a radioactively labeled substrate. In this manner, reaction products may be separated by chromatographic means, such as thin layer chromatography, HPLC or the like, and the relative amounts of the materials determined by scintillation counting.

While the foregoing approach has been found to work well by the inventors, there is no reason why other approaches might be employed, so long as one is able to determine whether a candidate substance has the ability to modify, alter or inhibit the enzyme being tested. Possible examples include spectrophotometric, gas chromatographic/ mass spectrophotometric or even using NMR analyses.

The invention is further described by reference to the drawings, a description of which is as follows:

Fig. 1 Expression cloning of steroid 5α-reductase. Female rat liver RNA wa size fractionated on 10-25% sucrose gradients and aliquots of RNA were assayed for steroid 5α-reductase activity in Xenopus oocytes. Peak activity fractions were used to construct an oriented cDNA library in a plasmid RNA expression vector. E. coli transformants from this library were pooled in groups of 150-200 clones and assayed for enzyme expression. A thin layer chromatography assay was employed in which the substrate testosterone (T) could be separated from androstenedione (A) and the 5α-reduced forms of these two steroids (DHT and 5αA, respectively). Sibling selection of a positive pool of clones was carried out as described in Example I.

Fig. 2. Dilution cloning of a liver steroid α-reductase cDNA. Xenopus oocytes were injected with RNA from the indicated source and assayed for steroid 5α-reductase activity by thin-layer chromatography

using [^{14}C]testosterone as a substrate as described in Example I. Lane 1, H_2O-injected; lane 2, RNA from female rat liver; lane 3, RNA synthesized in vitro from a pool of 150-200 cDNA clones; lane 4, RNA synthesized from cDNAs inoculated in a 96-well microtiter plate; lane 5, RNA synthesized from a pool of 12 clones corresponding to a row from the microtiter late; lane 6, RNA synthesized from eight clones corresponding to a column from this plate; and lane 7, RNA derived from a cDNA clone corresponding to the intersection of the row and column.

Chromatograms from the various experiments were exposed to Kodak XAR-5 film for 16 hours. In the chromatographic system employed, hydrophobic steroids migrate further than hydrophilic steroids. The positions of authentic steroid standards are shown on the left of the autoradiograms. T, testosterone, A, androstanedione, DHT, 5α-dihydrotestosterone, 5αA, 5α-androstanedione. An endogenous Xenopus enzyme in the oocytes converts testosterone into androstenedione. Steroids marked with an asterisk are uncharacterized metabolites derived from the 5α-reduced compounds by endogenous Xenopus enzymes (see Fig. 3). The amount of 5α-reduced metabolites in a given experiment varied depending on the batch of oocytes injected and is thus not calculated here.

Fig. 3. Substrate specificity of the cloned steroid 5α-reductase. Xenopus oocytes obtained from a single animal were injected with in vitro synthesized RNA derived from the steroid 5α-reductase cDNA clone and then assayed for enzyme activity using the indicated ^{14}C-labeled steroid substrates (5μM) in the absence (-) or presence (+) of the competitive inhibitor 4-MA (5 μM). The various steroids and metabolites are identified on the left and right of the autoradiograms: P, progesterone; 5αP, 5α-dihydroprogesterone; others are as indicated in the legend to Fig. 2. The amount of 5α-reduced metabolites for each substrate is indicated at the bottom of the figure and was determined by liquid scintillation counting after cutting out appropriate zones from the chromatograms. In lane 5 and 6, all radioactive derivatives of dihydrotestosterone were counted. In experiments not shown, the pattern of metabolites obtained when dihydrotestosterone was employed as a substrate was identical in both H_2O-injected and steroid 5α-reductase RNA-injected oocytes.

Fig. 4. Nucleotide sequence of the cDNA corresponding to the rat steroid 5α-reductase mRNA, predicted amino acid sequence, and hydropathy profile of the protein. A, nucleotides are numbered on the right-hand side. The amino acids are numbered above the sequence with position 1 arbitrarily assigned to the first methionine codon in the nucleotide sequence. Two polyadenylation signals are overlined. B, the sequence of the steroid 5α-reductase protein was subjected to a hydropathy analysis using the algorithm of Kyte and Doolittle (50). Sequences above the central dividing line are hydrophilic, and those below the line are hydrophobic.

Fig. 5. In vitro translation analysis of steroid 5α-reductase RNA. In vitro synthesized steroid 5α-reductase RNA was translated in a reticulocyte lysate as described in Example I. Additions to individual tubes are indicated above the autoradiogram. Approximately 8% of each translation reaction was analyzed by electrophoresis on 7-15% gradient polyacrylamide-sodium dodecyl sulfate gels. Size standards are indicated on the left. The band at M_r 45,000 represents an endogenous methionine binding protein in the reticulocyte lysate. The band corresponding to steroid 5α-reductase is indicated on the right of the autoradiogram.

Fig. 6. Characterization of the 5' and 3' ends of the steroid 5α-reductase cDNA and mRNA. A, expression of 3'-truncated RNAs in Xenopus oocytes. The steroid 5α-reductase cDNA plasmid was linearized with the indicated restriction enzyme and the resulting template was used to synthesize RNA in vitro. Oocytes were injected with the RNA and assayed for activity using testosterone as a substrate. The amount of 5α-reduced steroid metabolites was determined as described in the legend to Fig. 3. The values shown are the average of two or three separate experiments for each RNA. B, primer extension analysis of the 5' end of liver steroid 5-α-reductase mRNA, Ten μg of poly(A^+) mRNA from the indicted source was subjected to primer extension analysis as described in Example I. Size standards (STDS) are indicated on the left of the autoradiogram. Exposure times at -70 • C with an intensifying screen were 13 hours for lanes 1, 3, and 4, and 1 hour for lane 2. nt, nucleotides.

Fig. 7. cDNA sequence and predicted amino acid sequence of human steroid 5α-reductase. Nucleotides are numbered on the right with dots placed below the sequence every tenth nucleotide. Amino acid residues are numbered above the protein sequence. A potential polyadenylation signal (AATAAA) is overlined.

Fig. 8. Sequences of the human and rat steroid 5α-reductase proteins. The amino acid sequences of the human and rat enzymes are aligned to indicate homology between the two proteins. Identical residues are boxed. The single letter amino acid code is used with residues numbered on the right of the lineup.

Fig. 9. Expression of rat and human steroid 5α-reductase cDNAs in transfected COS cells. On day 2 after transfection, [^{14}C]testosterone was added to the medium to a final concentration of 2.5 μM. At the

indicated times, medium was removed from duplicate dishes and extracted with dichloromethane. Steroids were subjected to thin layer chromatography and scintillation counting as described in Materials and Methods.

Fig. 10. Inhibition of human steroid 5α-reductase activity in vitro by 4-MA and MK-906. COS cells were transfected with a human cDNA expression plasmid, lysed with a Polytron, and cell homogenates were assayed in vitro for steroid 5α-reductase activity as described in Materials and Methods. 40 μg of cellular protein was assayed in the presence of the indicated concentrations of 4-MA (Panel A) or MK-906 (Panel B) and 2 or 4 μM [^{14}C]testosterone. The data were plotted using an Apple IIe program. In panels A and B, the intersection of the two lines defines the K_i for the respective inhibitor (56).

Fig. 11. Inhibition of steroid 5α-reductase in transfected COS cells. COS cells were transfected on day 0 with an expression plasmid containing the human or rat steroid 5α-reductase cDNA. On day 2, a mixture consisting of 1 μM [^{14}C]testosterone and inhibitor (4-MA, top panel; MK-906, bottom panel) at the indicted concentration was added in ethanol to the medium of duplicate dishes. Conversion of testosterone into 5α-reduced products was monitored as described in Example II.

1. Introduction

A unique aspect of male sexual development is the requirement that the testis-derived hormone, testosterone[1], must be converted into dihydrotestosterone in target tissues that differentiate to form the male external genitalia (2). This conversion is catalyzed by a microsomal enzyme, steroid 5α-reductase, in the anlage of the external genitalia (1). The absence of steroid 5α-reductase activity underlies a rare form of male pseudohermaphroditism, originally termed pseudovaginal perineoscrotal hypospadias, in which the male external genitalia differentiate as female structures (13,14). In addition to its location in androgen-responsive tissues, high levels of steroid 5α-reductase activity are found in female rat liver but not in male rat liver (15). Whether this liver enzyme is the same as that in male target tissues is both controversial and unknown. Moreover, the factors that regulate the expression of this enzyme and the reason for the required conversion of testosterone into dihydrotestosterone for normal male differentiation are poorly understood (1).

The action of steroid 5α-reductase is a late event in male sexual development, a complex process that requires the correct developmental interpretation of both genetic and hormonal signals (2,16). The process is thought to begin after the establishment of chromosomal sex at fertilization with the expression of a master regulatory protein termed the testis determining factor. The gene for this protein has recently been cloned and shown to encode a protein with a structural motif (the zinc finger) commonly found in transcription factors (17,18). This finding is consistent with a role for this protein in the induction of a differentiation program leading to development of the testis (17). The testis in turn produces two hormones, testosterone, and a protein called the Mullerian inhibiting substance (19). The latter hormone causes regression of the Mullerian ducts, which are the anlage of female internal genitalia (2,19). Testosterone promotes development of the male internal genitalia (epididymides, vasa deferentia, and seminal vesicles), and after conversion to dihydrotestosterone by steroid 5α-reductase, the differentiation of the external male structures (penis, scrotum, and prostate) (2).

The actions of both testosterone and dihydrotestosterone in male development are mediated through a single protein, the androgen receptor, a recently cloned member of the steroid hormone receptor family (20-23). Genetic defects in the androgen receptor prevent the differentiation of both internal and external male structures (24). Although dihydrotestosterone has been shown to bind to the androgen receptor with higher affinity than testosterone (25), it is presently not known why the action of the receptor in promoting the differentiation of the external male genitalia requires synthesis of the higher affinity ligand. This requirement must be attributable to the presence of other regulatory factors in the development of the external genitalia (1).

Although the role of steroid 5α-reductase in male sexual differentiation has been elucidated, molecular insights into the gene and protein have not yet been possible due to the lack of genetic and immunochemical probes. The enzyme has been partially purified from the rat and shown to be an integral membrane protein of the endoplasmic reticulum or nuclear membrane (26). Much controversy exists in the literature as to the number of steroid 5α-reductase isozymes present in the liver and prostate of the rat and

[1] The abbreviations and trivial names used are: testosterone, 17β-hydroxy-4-androsten-3-one; dihydrotestosterone, 17β-hydroxy-5α-androstan-3-one; 4-MA, 17β-N,N-diethylcarbamoyl-4-methyl-4-aza-5α-androstan-3-one; androstenedione, androst-4-ene-3,17-dione; androstanedione, 5α-androstane-3,17-dione; progesterone, 4-pregnene-3,20-dione; 5α-dihydroprogesterone, 5α-pregnane-3,20-dione; kb, kilobase(s).

their cofactor requirements (27,28). Both the liver and prostate enzymes are inhibited in a competitive fashion by the steroid analogue 17β-N,N-diethylcarbamoyl-4-methyl-4-aza-5α-androstan-3-one (4-MA), suggesting that these proteins must at least share sequence homology in their substrate binding domains (10). Consistent with this prediction are the findings that the liver and prostate enzymes catalyse the reduction of similar steroid substrates, including testosterone, androstenedione, and progesterone (15,29).

The present disclosure specifically describes the cloning and sequence of DNA segments encoding a variety of steroid 5α-reductases, including those from rat liver, rat prostate and even from human sources. With these disclosures in light of the teachings herein, it is submitted that those of skill in the art will be enabled to prepare DNA segments encoding steroid 5α-reductases from any source desired without an undue amount of experimentation. Also disclosed are methods for employing these DNA segments to produce functional and assayable steroid 5α-reductases, which can be employed in a variety of manners, such as in the development of screening assays to identify inhibitors of the enzyme. It is also contemplated that these DNA segments can be employed in other manners, including, e.g., as probes for the identification of individuals who might carry defective steroid 5α-reductase genes, or certain alleles of this gene which predispose an individual to male pattern baldness, acne, hirsutism, and cancer of the prostate, or even other poorly described endocrine disorders of androgen metabolism.

2. Screening Assays

An important aspect of the invention is the use of recombinantly produced steroid 5α-reductase in screening assays for the identification of substances which may inhibit or otherwise modify or alter the enzymatic function of the enzyme. The use of recombinantly produced enzyme is of particular benefit because the naturally occurring enzyme is present in only small quantities and has proven difficult to purify. Moreover, this allows one a ready source of the human enzyme which has heretofore been lacking. The inventors have surprisingly discovered that the human enzyme is quite different from the steroid 5α-reductase obtained from species such as rat in terms of its sensitivity to various candidate substances. The importance of this is quite significant in that it indicates that where one seeks to identify a compound, e.g., that may function to inhibit the enzyme in man, that one should employ human species of steroid 5α-reductase for the screening assay. Furthermore, it suggests that previous studies where species other than the human enzyme were employed may not be accurate with respect to man.

The screening assays of the invention, in preferred embodiments, conveniently employ the enzyme directly from the recombinant host in which it is produced. This is achieved most preferrably by simply expressing the selected enzyme within the recombinant host, here a eukaryotic host, followed by preparing a crude homogenate which includes the enzyme. A portion of the crude homogenate is then admixed with an appropriate substrate of the enzyme, e.g., testosterone, progesterone, or androstenedione, along with the candidate substance to be tested. By comparing the action of the enzyme on the selected substrate in the presence or absence of the candidate substance, one can obtain information regarding the ability of the candidate substance to affect the activity of the enzyme.

In that most such screening assays in accordance with the invention will be designed to identify agents useful in inhibiting the conversion of testosterone, preferred assays will employ testosterone as the normal substrate.

There are believed to be a wide variety of embodiments which can be employed to determine the effect of the candidate substance on the enzymes of the invention, and the invention is not intended to be limited to any one such method. However, it will generally be desireable to employ a system wherein one can measure the ability of the enzyme to convert the subtrate employed to a particular product. One method employed by the inventors uses a labeled subtrate, which has been labeled in a manner such that the label is quantitatively retained in the resultant product. A convenient approach is the use of a radioactive label, such as C^{14} or H^3, which may be directly quantitated in both the substrate and the resultant product.

In preferred assays, the admixture containing the enzyme, substrate and candidate substance is allowed to incubate for a selected amount of time, and the resultant incubated mixture subjected to a separation means in order to separate the substrate remaining in the admixture from any product produced. Then, one simply measures the amount of each, e.g., versus a control to which no candidate substance has been added. This measurement can be made at various time points where velocity data is desired. From this, one may determine the ability of the candidate substance to alter or modify the function of the enzyme.

Numerous techniques are known which could be employed for the separation of the substrate from product, and all such methods are intended to fall within the scope of the invention. The inventors prefer to use thin layer chromatographic methods (TLC), as TLC-based methods are quick, accurate, inexpensive

and quite sensitive. However, other useful techniques might include, e.g., or other techniques such as HPLC, spectrophotometric, gas chromatographic/ mass spectrophotometric or even using NMR analyses. It is contemplated that any such technique may be employed so long as it is capable of differentiating between the enzyme substrate and product, and can be used to determine enzymatic function such as by identifying or quantifying the substrate and product.

3. Nucleic Acid Hybridization Embodiments

As mentioned, in certain aspects, the DNA sequence information provided by the invention allows for the preparation of relatively short DNA (or RNA) sequences having the ability to specifically hybridize to gene sequences of the selected steroid 5α-reductase gene. In these aspects, nucleic acid probes of an appropriate length are prepared based on a consideration of the selected steroid 5α-reductase gene sequence, e.g., a sequence such as that shown in Figure 7. The ability of such nucleic acid probes to specifically hybridize to the steroid 5α-reductase gene sequences lend them particular utility in a variety of embodiments. Most importantly, the probes can be used in a variety of assays for detecting the presence of complementary sequences in a given sample. However, either uses are envisioned, including the use of the sequence information for the preparation of mutant species primers, or primers for use in preparing other genetic constructions.

To provide certain of the advantages in accordance with the invention, the preferred nucleic acid sequence employed for hybridization studies or assays includes sequences that are complementary to at least a 10 to 30 or so long nucleotide stretch of the steroid 5α-reductase sequence, such as that shown in Figures 7. A size of at least 10 nucleotides in length helps to ensure that the fragment will be of sufficient length to form a duplex molecule that is both stable and selective. Molecules having complementary sequences over stretches greater than 10 bases in length ore generally preferred, though, in order to increase stability and selectivity of the hybrid, and thereby improve the quality and degree of specific hybrid molecules obtained. One will generally prefer to design nucleic acid molecules having gene-complementary stretches of 15 to 20 nucleotides, or even longer where desired. Such fragments may be readily prepared by, for example, directly synthesizing the fragment by chemical means, by application of nucleic acid reproduction technology, such as the PCR technology of U.S. Patent 4,603,102, or by introducing selected sequences into recombinant vectors for recombinant production.

Accordingly, the nucleotide sequences of the invention may be used for their ability to selectively form duplex molecules with complementary stretches of the gene. Depending on the application envisioned, one will desire to employ varying conditions of hybridization to achieve varying degree of selectivity of the probe toward the target sequence. For applications requiring a high degree of selectivity, one will typically desire to employ relatively stringent conditions to form the hybrids, for example, one will select relatively low salt and\or high temperature conditions, such as provided by 0.02M-0.15M NaCl at temperatures of 50 • C to 70 • C. These conditions are particularly selective, and tolerate little, if any, mismatch between the probe and the template or target strand.

Of course, for some applications, for example, where one desires to prepare mutants employing a mutant primer strand hybridized to an underlying template or where one seeks to isolate steroid 5α-reductase coding sequences for related species, functional equivalents, or the like, less stringent hybridization conditions will typically be needed in order to allow formation of the heteroduplex. In these circumstances, one may desire to employ conditions such as 0.15M-0.9M salt, at temperatures ranging from 20 • C to 55 • C. Cross-hybridizing species can thereby be readily identified as positively hybridizing signals with respect to control hybridizations. In any case, it is generally appreciated that conditions can be rendered more stringent by the addition of increasing amounts of formamide, which serves to destabilize the hybrid duplex in the same manner as increased temperature. Thus, hybridization conditions can be readily manipulated, and thus will generally be a method of choice depending on the desired results.

In certain embodiments, it will be advantageous to employ nucleic acid sequences of the present invention in combination with an appropriate means, such as a label, for determining hybridization. A wide variety of appropriate indicator means are known in the art, including radioactive, enzymatic or other ligands, such as avidin/biotin, which are capable of giving a detectable signal. In preferred embodiments, one will likely desire to employ an enzyme tag such a urease, alkaline phosphatase or peroxidase, instead of radioactive or other environmental undesirable reagents. In the case of enzyme tags, colorimetric indicator substrates are known which can be employed to provide a means visible to the human eye or spectrophotometrically, to identify specific hybridization with complementary nucleic acid-containing samples.

In general, it is envisioned that the hybridization probes described herein will be useful both as reagents in solution hybridization as well as in embodiments employing a solid phase. In embodiments involving a solid phase, the test DNA (or RNA) is adsorbed or otherwise affixed to a selected matrix or surface. This fixed, single-stranded nucleic acid is then subjected to specific hybridization with selected probes under desired conditions. The selected conditions will depend n the particular circumstances based on the particular criteria required (depending, for example, on the G+C contents, type of target nucleic acid, source of nucleic acid, size of hybridization probe, etc.). Following washing of the hybridized surface so as to remove nonspecifically bound probe molecules, specific hybridization is detected, or even quantified, by means of the label.

4. Biological Functional Equivalent Amino Acids

As noted above, it is believed that, where desired, modification and changes may be made in the structure of the steroid 5α-reductase and still obtain a molecule having like or otherwise desirable characteristics.

For example, certain amino acids may be substituted for other amino acids in a protein structure without appreciable loss of interactive binding capacity with structures such as antigen-binding regions of antibodies (or, e.g., binding sites on substrate molecules. Since it is the interactive capacity and nature of a protein that defines that protein's biological functional activity, certain amino acid sequence substitutions can be made in a protein sequence (or, of course, its underlying DNA coding sequence) and nevertheless obtain a protein with like or even counterveiling properties (e.g., antagonistic v. agonistic). It is thus contemplated by the inventors that various changes may be made in the sequence of the peptides (or underlying DNA) without appreciable loss of their biological utility or activity.

The importance of the hydropathic index of amino acids in conferring interactive biologic function on a protein has been discussed generally by Kyte et al. (50), wherein it is found that certain amino acids may be substituted for other amino acids having a similar hydropathic index or score and still retain a similar biological activity. As displayed in the table below, amino acids are assigned a hydropathic index on the basis of their hydrophobicity and charge characteristics. It is believed that the relative hydropathic character of the amino acid determines the secondary structure of the resultant protein, which in turn defines the interaction of the protein with substrate molecules.

TABLE I

AMINO ACID	HYDROPATHIC INDEX
Isoleucine	4.5
Valine	4.2
Leucine	3.8
Phenylalanine	2.8
Cysteine/cystine	2.5
Methionine	1.9
Alanine	1.8
Glycine	-0.4
Threonine	-0.7
Tryptophan	-0.9
Serine	-0.8
Tyrosine	-1.3
Proline	-1.6
Histidine	-3.2
Glutamic Acid	-3.5
Glutamine	-3.5
Aspartic Acid	-3.5
Asparagine	-3.5
Lysine	-3.9
Arginine	-4.5

It is proposed that where an amino acid has a hydropathic index of within ±2 that of the base amino acid, and more preferably within ±1, such a change should nevertheless provide a protein having a similar,

and perhaps even improved, functional activity. Thus, for example, it is proposed the isoleucine, which has a hydrophatic index of +4.5, can be substituted for valine (+4.2) or leucine (+3.8), and still obtain a protein having similar biologic activity. Alternatively, at the other end of the scale, it is proposed that lysine (-3.9) can be substituted for arginine (-4.5), and so on.

Accordingly, these amino acid substitutions are generally based on the relative similarity of R-group substituents, for example, in terms of size, electrophilic character, charge, and the like. In general, exemplary substitutions which take various of the foregoing characteristics into consideration include the following:

TABLE II

Original Residue	Exemplary Substitutions
Ala	gly; ser
Arg	lys
Asn	gln; his
Asp	glu
Cys	ser
Gln	asn
Glu	asp
Gly	ala
His	asn; gln
Ile	leu; val
Leu	ile; val
Lys	arg
Met	met; leu; tyr
Ser	thr
Thr	ser
Trp	tyr
Tyr	trp; phe
Val	ile; leu

5. Site-Specific Mutagenesis

Site-specific mutagenesis is a technique useful in the preparation of second generation proteins, or biologically functional equivalent proteins or peptides, derived from the sequences thereof, through specific mutagenesis of the underlying DNA. The technique further provides a ready ability to prepare and test sequence variants, for example, incorporating one or more of the foregoing considerations, by introducing one or more nucleotide sequence changes into the DNA. Site-specific mutagenesis allows the production of mutants through the use of specific oligonucleotide sequences which encode the DNA sequence of the desired mutation, as well as a sufficient number of adjacent nucleotides, to provide a primer sequence of sufficient size and sequence complexity to form a stable duplex on both sides of the deletion junction being traversed. Typically, a primer of about 17 to 25 nucleotides in length is preferred, with about 5 to 10 residues on both sides of the junction of the sequence being altered.

In general, the technique of site-specific mutagenesis is well known in the art, as exemplified by publications such as reference (61), incorporated herein by reference. As will be appreciated, the technique typically employs a phage vector which exist in both a single stranded and double stranded form. Typical vectors useful in site-directed mutagenesis include vectors such as the M13 phage, for example, as disclosed by reference 62, incorporated hereby in reference. These phage are readily commercially available and their use is generally well known to those skilled in the art.

In general, site-directed mutagenesis in accordance herewith is performed by first obtaining a single-stranded vector which includes within its sequence a DNA sequence which encodes all or a portion of the steroid 5α-reductase sequence. An oligonucleotide primer bearing the desired mutated sequence is prepared, generally synthetically, for example, by the method of reference 63. This primer is then annealed with the singled-stranded vector, and subjected to DNA polymerizing enzymes such as E. coli polymerase I Klenow fragment, in order to complete the synthesis of the mutation-bearing strand. Thus, a heteroduplex is formed wherein one strand encodes the original non-mutated sequence and the second strand bears the

desired mutation. This heteroduplex vector is then used to transform appropriate cells such as E. coli cells and clones are selected which include recombinant vectors bearing the mutated sequence arrangement.

6. Host Cell Cultures and Vectors

In general, of course, prokaryotes are preferred for the initial cloning of DNA sequences and constructing the vectors useful in the invention. For example, E. coli K12 strains may be particularly useful. Other microbial strains which may be used include E. coli B, and E. coli X 1776 (ATCC No. 31537). These examples are, of course, intended to be illustrative rather than limiting.

Prokaryotes may also be used for expression. The aforementioned strains, as well as E. coli W3110 (F-, lambda-, prototrophic, aTCC No. 273325), bacilli such as Bacillus subtilus, or other enterobacteriacea such as Salmonella typhimurium or Serratus marcesans, and various Pseudomonas species may be used.

In general, plasmid vectors containing replicon and control sequences which are derived from species compatible with the host cell are used in connection with these hosts. The vector ordinarily carries a replication site, as well as marking sequences which are capable of providing phenotypic selection in transformed cells. For example, E. coli is typically transformed using pBR 322, a plasmid derived from an E. coli species (see, e.g., Reference 64). pBR 322 contains genes for ampicillin and tetracycline resistance and thus provides easy means for identifying transformed cells. The pBR plasmid, or other microbial plasmid or phage must also contain, or be modified to contain, promoters which can be used by the microbial organism for expression of its own proteins.

Those promoters most commonly used in recombinant DNA construction include the B-lactamase (penicillinase) and lactose promoter systems (65-67) and a tryptophan (TRP) promoter system (68-69). While these are the most commonly used, other microbial promoters have been discovered and utilized, and details concerning their nucleotide sequences have been published, enabling a skilled worker to ligate them functionally with plasmid vectors (70).

In addition to prokaryotes, eukaryotic microbes, such as yeast cultures may also be used. Saccharomyces cerevisiase, or common baker's yeast is the most commonly used among eukaryotic microorganisms, although a number of other strains are commonly available. For expression in Saccharomyces, the plasmid YRp7, for example, is commonly used (71-73). This plasmid already contains the trpl gene which provides a selection marker for a mutant strain of yeast lacking the ability to grow in tryptophan, for example ATCC No. 44076 or PEP4-1 (74). The presence of the trpl lesion as a characteristic of the yeast host cell genome then provides an effective environment for detecting transformation by growth in the absence of tryptophan.

Suitable promoting sequences in yeast vectors include the promoters for 3-phosphoglycerate kinase (75) or other glycolytic enzymes (76,77), such as enolase, glyceraldehyde-3-phosphate dehydrogenase, hexokinase, pyruvate decarboxylase, phosphofructokinase, glucose-6-phosphate isomerase, 3-phosphoglycerate mutase, pyruvate kinase, triosephosphate isomerase, phosphoglucose isomerase, and glucokinase. In constructing suitable expression plasmids, the termination sequences associated with these genes are also ligated into the expression vector 3' of the sequences desired to be expressed to provide polyadenylation of the mRNA and termination. Other promoters, which have the additional advantage of transcription controlled by growth conditions are the promoter region for alcohol dehydrogenase 2, isocytochrome C, acid phosphatase, degradative enzymes associated with nitrogen metabolism, and the aforementioned glyceraldehyde-3-phosphate dehydrogenase, and enzymes responsible for maltose and galactose utilization. Any plasmid vector containing a yeast-compatible promoter, origin or replication and termination sequences is suitable.

In addition to microorganisms, cultures of cells derived from multicellular organisms may also be used as hosts. In principle, any such cell culture is workable, whether from vertebrate or invertebrate culture. However, interest has been greatest in vertebrate cells, and propogation of vertebrate cells in culture (tissue culture) has become a routine procedure in recent years (78). Examples of such useful host bell lines are AtT-20 VERO and HeLa cells, Chinese hamster ovary (CHO) cell lines, and W138, BHK, COS-7 293 and MDCK cell lines. Expression vectors for such cells ordinarily include (if necessary) an origin of replication, a promoter located in front of the gene to be expressed, along with any necessary ribosome binding sites, RNA splice sites, polyadenylation site, and transcriptional terminator sequences.

For use in mammalian cells, the control functions on the expression vectors are often provided by viral material. For example, commonly used promoters are derived from polyoma, Adenovirus 2, Cytomegalovirus and most frequently simian Virus 40 (SV40). The early and late promoters of SV40 virus are particularly useful because both are obtained easily from the virus as a fragment which also contains the SV40 viral origin of replication (79). Smaller or larger SV40 fragments may also be used, provided there

is included the approximately 250 bp sequence extending from the Hind III site toward the Bgl I site located in the viral origin or replication. Further, it is also possible, and often desirable, to utilize promoter or control sequences normally associated with the desired gene sequence, provided such control sequences are compatible with the host cell systems.

An origin of replication may be provided with by construction of the vector to include an exogenous origin, such as may be derived from SV40 or other viral (e.g., Polyoma, Adeno, VSV, BPV, CMV source, or may be provided by the host cell chromosomal replication mechanism. If the vector is integrated into the host cell chromosome, the latter is often sufficient.

7. pCMV Eukaryotic Expression Vectors

The pCMV plasmids are a series of mammalian expression vectors constructed by individuals in the Department of Molecular Genetics, University of Texas Southwestern Medical Center. The vectors are designed for use in essentially all cultured cells and work extremely well in SV40-transformed simian COS cell lines. The pCMV1, 2, 3, and 5 vectors differ from each other in certain unique restriction sites shown in the polylinker region above each plasmid. The pCMV4 vector differs from these 4 plasmids in containing a translation enhancer in the sequence prior to the polylinker.

The universal components of the pCMV plasmids are as follows. The vector backbone is pTZ18R (Pharmacia), and contains a bacteriophage f1 origin or replication for production of single stranded DNA and an ampicillin-resistance gene. The CMV region consists of nucleotides -760 to +3 of the powerful promote-regulatory region of the human cytomegalovirus (Towne stain) major immediate early gene (80).

The polylinker region may be synthesized on an Applied Biosystem's machine. The human growth hormone fragment (hGH) contains transcription termination and poly-adenylation signals representing sequences 1533 to 2157 of this gene (82). There is an Alu middle repetitive DNA sequence in this fragment. Finally, the SV40 origin of replication and early region promoter-enhancer (white box) was derived from the pcD-X plasmid (HindII to PstI fragment) described in (83). The promoter in this fragment is oriented such that transcription proceeds away from the CMV/hGH expression cassette.

The pCMV plasmids are distinguished from each other by which restriction enzyme sites are unique in the polylinker and by the presence or absence of the translation enhancer. The starting pCMV1 plasmid has been progressively modified to render unique an increasing number of sites in the polylinker. To create pCMV2, one of two EcoRI sites in pCMV1 were destroyed. To create pCMV3, pCMV1 was modified by deleting a short segment from the SV40 region (StuI to ^R EcoRI), and in so doing made unique the PstI, SalI, and BamHI sites in the polylinker. To create pCMV4, a synthetic fragment of DNA corresponding to the 5'-untranslated region of a mRNA transcribed from the CMV promote. The sequence acts as a transnational enhancer by decreasing the requirements for initiation factors in protein synthesis (81). To create pCMV5, a segment of DNA (HpaI to EcoRI) was deleted from the SV40 origin region of pCMV1 to render unique all sites in the starting polylinker.

The pCMV vectors have been employed in simian COS cells, mouse L cells, CHO cells, and HeLa cells. In several side by side comparisons they have yielded 5- to 10-fold higher expression levels in COS cells than SV40-based vectors. The pCMV vectors have been used to express the LDL receptor, nuclear factor 1, G_s alpha protein, protein phosphatase, synaptophysin, synapsin, insulin receptor, flu hemmagglutinin, androgen receptor, sterol 26-hydroxylase, steroid 17- and 21-hydroxylase, cytochrome P-450 oxidoreductase, beta-adrenergic receptor, folate receptor, cholesterol side chain cleavage enzyme, and a host of other cDNAs. It should be noted that the SV40 promoter in these plasmids can be used to express other genes such as dominant selectable markers. Finally, there is an ATG sequence in the polylinker between the HindIII and PstI sites that may cause spurious translation initiation. Avoid this codon if possible in your expression plasmids. A paper describing the construction and use of pCMV1 and pCMV4 has been published (52).

8. Examples

Examples have been included in order to illustrate preferred modes of the invention. Certain aspects of the following examples are described in terms of techniques and procedures found or contemplated by the present inventors to work well in the practice of the invention. These examples are exemplified through the use of standard laboratory practices of the inventor. In light of the present disclosure and the general level of skill in the art, those of skill will appreciate that the following examples are intended to be exemplary only and that numerous changes, modifications and alterations can be employed without departing from the spirit and scope of the invention.

EXAMPLE I

Expression Cloning of Rat Steroid 5α-Reductas

This example describes techniques employed by the inventors for the cloning of cDNAs encoding the rat liver and prostate steroid 5α-reductase enzymes. Because there was no known sequence information for this enzyme upon which to base the construction of oligonucleotide probes, a novel approach was developed, based in part on a strategy employing Xenopus oocyte expression cloning employed in the isolation of lymphokines (30), neurotransmitter receptors (31-33), and membrane transporters (34). As will be seen from the results discussed below, the amino acid sequences deduced from the cDNAs reveal that the liver and prostate forms of steroid 5α-reductase are identical. However, RNA blotting experiments suggested that the expression of steroid 5α-reductase in these two tissues is differentially regulated by testosterone.

A. Protocols Employed

1. Steroid 5α-Reductase Enzyme Assay

Stage 5 and 6 oocytes were surgically removed from female Xenopus laevis (NASCO, Fort Atkinson, WI) and collagenase-treated as described by Julius et al. (33). Oocytes were injected with 50-100 nl of RNA (1 μg/μl) as described by Peacock et al. (35). After injection the oocytes were incubated at 19 °C for 24 hours in modified Barth's saline solution (35) containing 1 mg/ml bovine serum albumin to allow expression of the injected RNA. Five to ten viable oocytes were then transferred to 1 ml of modified Barth's saline solution containing 5 μM ^{14}C-labeled steroid (50 mCi/mmol, Du Pont-New England Nuclear), and incubated at 37 °C for 2-24 h. This temperature-jump protocol is based on the observation that expression of mRNA in Xenopus is maximal a 19 °C, whereas rat steroid 5α-reductase expressed in Xenopus has a temperature optima of 37 °C. After the 37 °C incubation, the oocytes were homogenized in the incubation medium and steroid was extracted with 10 ml of dichloromethane. The solvent was evaporated under air and the residue was dissolved in 0.1 ml of chloroform/methanol (2:1, v/v) and subjected to thin-layer chromatography using Siica Gel 60 thin-layer chromatography plates (E. merck, 5748-7, Darmstadt, West Germany). The chromatoplates were autoradiographed for 18 hours at -70 °C and the radioactive zones were cut out and subjected to liquid scintillation counting in Complete Counting Cocktail (Research Products International). The identities of the products were determined by comparison to the R_F values of known standards.

2. cDNA Cloning

Total RNA from female rat liver was extracted by a guanidinium isothiocyanate/CsC1 procedure (36). Poly(A^+)-enriched RNA was isolated and size-fractionated by density gradient centrifugation on 10-25% (w/v) sucrose gradients containing methylmercury hydroxide (37). After centrifugation at 4 °C for 15 hours at 76,800 X g, aliquots of RNA from each gradient fraction were assayed for steroid 5α-reductase mRNA by injection into Xenopus oocytes. Positive fraction from the sucrose gradients were combined and the RNA was concentrated by ethanol precipitation. First strand cDNA was synthesized using mRNA pretreated with 2.5 mM methylmercury hydroxide and AGCGGCCGC$(T)_{20}$ as a primer. Second strand synthesis, EcoRI methylation, flushing of ends with bacteriophage T4 DNA polymerase, and addition of phosphorylated EcoRI linker were performed according to standard procedures (36). The resulting cDNA was digested with NotI and EcoRI and size-fractionated on a 1% (w/v) agarose gel. Complementary DNAs greater than 1.3 kb were inserted into the EcoRI and NotI sites of Bluescript (Stratagene, LA, Jolla, CA). Recombinant plasmids were propagated in Escherichia coli DH5αF'IQ (GIBCO). A rat ventral prostate cDNA library was constructed as described above except that random hexanucleotides were used as primers and total poly(A^+) RNA was used as template. Size-fractionated cDNAs derived from prostate mRNA were inserted into the EcoRI site of λZapII (Stratagene). Recombinant bacteriophage were propagated in E. coli XL1-Blue. Bluescript plasmids were subsequently rescued from λZap recombinants by superinfection with helper F1 bacteriophage.

In the initial screening of the female rat liver cDNA library, plasmids minipreps were prepared from 20 pools containing 150-200 cDNA clones/pool. Plasmid DNA was linearized with NotI and RNA was transcribed in vitro using bacteriophage T7 RNA polymerase (Pharmacia LKB Biotechnology Inc.) as described by Julius et al. (33). Xenopus oocyte injection was carried out as described above. Plasmid DNA from one positive pool was retransformed and 960 colonies were randomly picked into individual 0.3-ml cultures maintained in 96-well microtiter plates. Plasmid DNAs were subsequently prepared from pools of

100-µl aliquots from each well and assayed by microinjection. Sibling selection from the microtiter plate was carried out by matrix analysis.

3. Nucleic Acid Sequencing and Primer Extension

Overlapping fragments from both DNA strands were subcloned into bacteriophage M13 vectors and sequenced by automated methods (38) using an Applied Biosystems model 370A DNA sequencer. For primer extension analysis, an antisense oligonucleotide complimentary to nucleotides 70-109 of Fig. 4A was annealed at 68 °C to rat liver poly(A^+) RNA and extended with reverse transcriptase as described by Sudhof et al. (39). Direct RNA sequencing of the steroid 5α-reductase mRNA was carried out as described by Geliebter et al. (40).

4. In Vitro Translation of RNA

Approximately 100 ng of RNA was translated in vitro using [^{35}S]methionine (1100 Ci/mmol) and a rabbit reticulocyte lysate (Promega, Madison, WI) in the presence or absence of dog pancreas microsomes (41). After incubation for 1 hour at 30 °C, the reactions were terminated by adding cycloheximide to a final concentration of 0.2 mM or RNase A to 2 mg ml^{-1}. Experiments with products translated in vitro in the presence of 50 µg/ml trypsin (GIBCO) were performed with or without 2% (w/v) Triton X-100 (Boehringer Mannheim) for 30 min at 22 °C. The protease reactions were terminated by adding soybean trypsin inhibitor (Cappel, Malvern, PA) to a final concentration of 1 mg ml^{-1}.

5. Physiology Experiments

Studies were designed to allow comparison of mRNA levels in liver and prostate of normal rats, of 7-day castrated animals, of 10-day castrated animals, and of normal or 10-day castrated animals given testosterone on days 7-9 of the experiment. Sexually mature Sprague-Dawley male rats were castrated by standard surgical procedures on day 0. On day 7, experimental groups were subcutaneously injected for 3 consecutive days with 2 mg of testosterone acetate or testosterone propionate dissolved in 0.2 ml of sesame oil (15). Control animals were injected with sesame oil alone. On day 10 of the experiment, RNA was prepared from the livers and prostates of up to 15 animals in each experimental group, and analyzed by blotting as described in the legend to Fig. 6.

B. Results and Discussion

1. Expression Cloning of the Rat Liver Steroid 5α-Reductase cDNA

Fig. 1 outlines the strategy used to obtain a full length cDNA for the rat liver steroid 5α-reductase. As a source of mRNA, female rat liver was used, which for physiologically unknown reasons expresses high levels of steroid 5α-reductase enzyme activity (26). Microinjection into Xenopus oocytes indicated that this mRNA could direct the synthesis of an enzyme that catalyzed the conversion of steroids into their 5α-reduced forms (see below). Sucrose gradient fractionation of rat liver mRNA indicated that this activity was encoded by an mRNA of about 2.5 kb (Fig. 1). Similar results have recently been reported by Farkash et al. (42). The mRNA in this fraction was converted into cDNA, size-fractionated, and cloned into an RNA expression vector. To avoid problems with anti-sense inhibition, the cDNA library was constructed in an oriented manner (Fig. 1). Twenty pools, each containing 150-200 cDNA clones, were then used to synthesize mRNA that was in turn injected into oocytes to allow determination of steroid 5α-reductase activity by thin-layer chromatography analysis. From one active pool, a near full length cDNA encoding this enzyme was subsequently isolated by dilution cloning (Fig. 1).

Fig. 2 illustrates the results of thin-layer chromatography assays from the dilution cloning. In all studies, assay of steroid 5α-reductase activity in injected oocytes was carried out using a temperature-jump protocol as detailed in section A.2. above. Microinjection of water into Xenopus oocytes revealed an endogenous activity capable of converting the testosterone substrate into androstenedione, and little or no ability to convert these steroids into their 5α-reduced forms (lane 1). In contrast, when female rat liver mRNA was injected, the oocytes expressed an activity that generated both dihydrotestosterone and 5α-androstanedione, as well as at least two other steroid metabolites (lane 2). These latter unidentified steroids were derived from the 5α-reduced metabolites generated by the injected mRNA (see below).

Lane 3 of Fig. 2 shows the results obtained when RNA was synthesized from one of the initial 20 cDNA plasmid pools that contained 150-200 independent clones. The spectrum of steroid metabolites observed was identical to that seen upon injection of liver mRNA, indicating that this pool must contain at least one steroid 5α-reductase cDNA. The cDNAs from this pool were retransformed into E. coli and individual colonies were picked into microtiter plates. Lane 4 shows the results obtained after microinjection of RNA prepared from plasmids isolated from a 96-well plate that contained a steroid 5α-reductase cDNA from this transformation. Subsequent analysis of mRNA from pools of plasmids corresponding to the rows and columns of this microtiter plate identified a row (lane 5) and column (lane 6) containing a steroid 5α-reductase plasmid. The intersection of this row and column on the microtiter plate localized the positive cDNA (lane 7).

2. Substrate Specificity of Cloned Liver steroid 5α-Reductase.

RNA synthesized from the steroid 5α-reductase cDNA plasmid identified in Fig. 2 was microinjected into oocytes and allowed to express for a 24-hour period. The oocytes were then incubated with different radiolabeled steroids for an additional 24 hours and the products formed were analyzed by thin-layer chromatography. Fig. 3, lane 1, shows the typical pattern of 5α-reduced metabolites formed from testosterone. Lane 2 indicated that co-incubation of the injected eggs with equimolar amounts of testosterone and the competitive steroid 5α-reductase inhibitor 4-MA resulted in a substantial decrease in the formation of these products. As a control for nonspecific inhibition, the conversion of testosterone into androstenedione catalyzed by an endogenous Xenopus enzyme (presumably a 17β-hydroxysteroid dehydrogenase) (43), was not inhibited by 4-MA in this experiment (lane 2). The data in lane 3 and 7 indicate that both androstenedione and progesterone were substrates for the cloned enzyme. As with testosterone, 4-MA efficiently blocked the reduction of these steroids (lanes 4 and 8, respectively). When radiolabeled dihydrotestosterone was used as a substrate (lane 5), the inhibitor had no effect on the conversion of this compound into other 5α-reduced metabolites by endogenous Xenopus enzymes (lane 6).

3. Sequences of Liver Steroid 5α-Reductase

Fig. 4A shows the nucleotide sequence of the liver steroid 5α-reductase cDNA and the deduced amino acid sequence of the protein. The cDNA insert in the expressing clone was 2,465 base pairs in length and included a long 3'-untranslated region of 1,691 base pairs and an extended translation reading frame of 765 base pairs. A potential polyadenylation signal is present at position 2,446, upstream of a tract of A residues, suggesting that the 3' end of this cDNA is authentic. In the predicted amino acid sequence, there are three methionine residues in the first 19 amino acids. The context of the first ATG is identical in six out of nine nucleotides with the ideal Kozak consensus sequence (44), suggesting that this codon may specify the amino-terminal methionine of steroid 5α-reductase. With this assumption, the open reading frame would encode a hydrophobic protein of 255 amino acids with a predicted M_r of 29,343. Over 50% of the amino acids in the protein sequence have hydrophobic side chains. Consistent with this amino acid composition, a hydropathy plot (Fig. 4B) suggests a protein with many hydrophobic regions. A comparison of the sequence shown in Fig. 4A to others in the National Biomedical Research Foundation protein data bank and the GenBank DNA sequence collection did not reveal any sequences that were homologous to steroid 5α-reductase.

4. Characterization of Steroid 5α-Reductase Protein and mRNA

Several reports in the literature have identified a rat liver protein of M_r 50,000 that either has steroid 5α-reductase activity or can be cross-linked to a photoactivatable derivative of 4-MA (45,46). To ensure that the sequence shown in Fig. 4A represented the complete coding region of steroid 5α-reductase, three kinds of studies were conducted. First, as shown in Fig. 5, in vitro translation in a rabbit reticulocyte lysate of RNA generated from the steroid 5α-reductase cDNA yielded a protein product with an apparent M_r of 26,000 (lane 3). When the translation reactions were carried out in the presence of dog pancreas microsomes, a protein product of identical size was observed (lane 4), suggesting the absence of a cleavable signal sequence in this protein. That the steroid 5α-reductase translated in vitro was incorporated into microsomes was demonstrated by protease protection experiments. If the vesicular structure of the microsomes was maintained, the translated product was largely resistant to digestion by trypsin (lane 5). However, if the microsomes were disrupted with the detergent Triton X-100 prior to protease treatment, then the steroid 5α-reductase protein was susceptible to digestion (lane 6).

The approximate location of the carboxyl terminus of the protein was next determined by analyzing the expression of RNA derived from a series of 3'-truncated derivatives of the cDNA. The steroid 5α-reductase cDNA plasmid was linearized by cleavage with four restriction enzymes that left intact or removed progressively large portions of the predicted 3'-untranslated region and/or carboxyl terminus of the protein. RNA was transcribed in vitro from these templates, microinjected into oocytes, and the oocytes were assayed for steroid 5α-reductase activity using testosterone as a substrate.

As summarized in Fig. 6A, expression of the intact steroid 5α-reductase RNA resulted in the reduction of 67% of the testosterone substrate. Removal of 1474 nucleotides from the 3'-untranslated region of the mRNA did not substantially affect expression of enzyme activity (BamHI-cleaved template, Fig. 6A). However, removal of 1830 nucleotides from the 3' end, which removes 47 amino acid residues from the predicted carboxyl terminus of the protein, eliminated steroid 5α-reductase activity (PvuII-cleaved template, Fig. 6A). Similar results were obtained with a truncated RNA that removed 57 residues from the carboxyl terminus of the protein (SacI-cleaved template, Fig. 6A). In experiments not shown, all of these mRNAs yielded a protein of the appropriate size after in vitro translation in a reticulocyte lysate.

The amino-terminal region of steroid 5α-reductase was examined by carrying out primer extension experiments on liver mRNA. An oligonucleotide primer 40 bases in length and complementary to nucleotides 70-109 of Fig. 4A was radiolabeled, annealed to mRNA from female and male rat liver, and extended with reverse transcriptase. As shown in Fig. 6B, a single product of 125 nucleotides was detected when RNA from female or male liver was used as a template. These results are consistent with a single 5' end for the steroid 5α-reductase mRNA in this tissue and suggests that the cDNA sequence shown in Fig. 4A represents a near full length clone. In additional experiments not shown, the steroid 5α-reductase mRNA in female rat liver was sequenced directly using the above primer. The results indicated that the mRNA extends only 17 nucleotides upstream of the 5' end of the cDNA sequence shown in Fig. 4A. There were no inframe translation stop codons in this 5' sequence.

4. The Liver and Ventral Prostate Forms of Steroid 5α-Reductase Are Identical

To determine if the steroid 5α-reductase activities in the liver and prostate were due to the expression of a single mRNA, a randomly primed cDNA library derived from ventral prostate mRNA was screened with the insert derived from the liver cDNA clone. A single prostate cDNA was isolated after screening approximately 150,000 independent clones. DNA sequence analysis of the 5' and 3' ends of this clone indicated that it began at nucleotide 1 and terminated at nucleotide 1955 of the liver cDNA sequence shown in Fig. 4A. The sequences were identical between the two clones in these regions. The complete coding region of the prostate-derived cDNA was further subjected to DNA sequence analysis and comparison to that of the liver cDNA again revealed no differences. These results suggested that the enzyme activities in these two tissues were the consequence of expression of the same mRNA.

5. Conclusions

The data presented here provide evidence that a single mRNA encodes the enzyme steroid 5α-reductase in both the liver and prostate of the rat. DNA sequence analysis indicates that this mRNA encodes a hydrophobic protein with a M_r of 29,000. This size was confirmed by in vitro translation, mRNA truncation experiments, and primer extension analysis of live mRNA. Blot hybridization analyses of RNA and genomic DNA support the existence of a single mRNA and gene for steroid 5α-reductase.

The contributions of steroid 5α-reductase and its product dihydrotestosterone to male sexual development are clearly illustrated by the clinical syndrome of pseudovaginal perineoscrotal hypospadias (24). In addition to this developmental role, dihydrotestosterone in the mature organism is involved in the normal maintenance of many different cellular and organ processes. In fact, the suggestion has been made that it is dihydrotestosterone and not testosterone that is the more important androgen in this regard (1). As such, abnormal expression of steroid 5α-reductase and subsequent dihydrotestosterone synthesis may contribute to a large number of human diseases and endocrine abnormalities.

Localized overproduction of dihydrotestosterone in the prostate is postulated to be a factor in benign prostate hypertrophy, a condition that affects a majority of elderly men (47). Similarly, dihydrotestosterone has been implicated in the formation of acne and in the manifestation of male pattern baldness (48). Finally, a role for this hormone in the development and/or susceptibility to cancer of the prostate, the second most prevalent form of cancer in the United States, has been hypothesized (49). The precise contribution of steroid 5α-reductase to these disease states has so far remained uncertain due to the absence of biochemical and genetic tools. The results presented here clarify many controversies that have existed in

the literature for 20 years concerning this enzyme and they may provide these necessary tools.

EXAMPLE II

Cloning and Expression of Human Steroid 5α-Reductase

The present example describes the methods employed by the inventors to prepare DNA segments encoding human steroid 5α-reductase, through the cloning of the human steroid 5α-reductase gene. In this procedure, the rat cDNA described in Example I was used as a cross-hybridization probe to screen a human prostate cDNA library. A 2.1 kilobase cDNA was identified and DNA sequence analysis indicated that the human steroid 5α-reductase was a hydrophobic protein of 259 amino acids with a predicated M_r of 29,462. However, a comparison of the human and rat protein sequences revealed only a 60% identity. Transfection of expression vectors containing the human and rat cDNAs into simian COS cells resulted in the synthesis of high levels of steroid 5α-reductase enzyme activity. Both enzymes expressed in COS cells showed similar substrate specificities for naturally occurring steroid hormones.

Through the use of screening assays developed for use in connection with the invention, however, the inventors surprisingly discovered that potential therapeutic inhibitors demonstrate marked differences in their effect on the human enzyme as compared to the rat species. For example, synthetic 4-aza-steroids demonstrated marked differences in their abilities to inhibit the human and rat steroid 5α-reductases.

A. Protocols Employed

1. Materials

Radiolabeled steroids were obtained from Du Pont-New England Nuclear and steroid standards were from Sigma and Steraloids, Inc. The 4-azasteroids, 4-MA (17β-N,N-diethylcarbamoyl-4-methyl-4-aza-5α-androtane-3-one) and MK-906 (17β-N-t-butylcarbamoyl-4-aza-5α-androst-l-en-3-one) were gifts of Merck Sharp and Dohme Research Laboratories. Inhibitors were subjected to chemical ionization-mass spectrometry to confirm their identity prior to use.

2. cDNA Cloning

Two cDNA libraries were constructed from human prostate mRNA. Fir the first, cDNA provided by Dr. M.J. McPhaul of The University of Texas Southwestern Medical Center (22) was ligated into the bacteriophage λgt10 vector as described in Example I. For the second, prostate tissue from a subject undergoing surgery for prostatic hyperplasia was obtained and used for the isolation of polyadenylated RNA (51). A size-fractionated cDNA library was subsequently prepared (see Example I) in λgt10 (51). Clones from these libraries were screened by using hybridization conditions of reduced stringency (51). DNA sequence analysis was carried out using automated methods on an Applied Biosystems (Foster City, CA) model 370A DNA Sequencer. RNA blotting was performed as described (51).

3. Expression Vector Construction.

A rat steroid 5α-reductase cDNA corresponding to nucleotides 1-1962 (see Example I) was ligated into the pCMV4 expression vector (52). A human cDNA corresponding to nucleotides 1 to 842 of Fig. 7 was initially ligated into pCMV4. To modify this poorly expressed human cDNA (see below), two oligonucleotides derived from the 5'-end of the cDNA (5' ATAGATCTACCATGGCAACGGCGA 3'), or from the 3'-untranslated region (5'AAAGTCCATAGAGAAGCGCCATTGG 3') were employed in a polymerase chain reaction (53) to alter the human cDNA as described below. After amplification, the product was ligated into pCMV4.

4. Expression of Steroid 5α-Reductase cDNAs in COS Cells.

Simian COS-M6 cells were transfected as described (52). The assay of steroid 5α-reductase activity in intact cells was carried out as described in Example I except that [^{14}C]-labeled steroid dissolved in ethanol were added to the transfected cell medium and subsequent organic extractions were carried out with dichloromethane. Thin layer chromatography and liquid scintillation counting were performed as described in Example I. To determine IC_{50} values for the 4-MA and MK-906 inhibitors, a mixture of [^{14}C]testosterone and inhibitor in ethanol was added to transfected cell medium, incubated at 37°C for 2 hr, and treated as

above.

To assay steroid 5α-reductase activity in vitro, cells were harvested 48 hours after transfection, washed once with phosphate buffered saline and either frozen in liquid N_2 or homogenized directly with a Polytron at a protein concentration of 2 mg/ml in 10 mM potassium phosphate (pH 7.4), 150 mM KC1 and 1 mM EDTA. A typical assay contained 10 to 50 μg of cell homogenate protein in 0.5 ml of 0.1 M potassium phosphate buffer (pH 6.6, rat enzyme; pH 7.0, human enzyme). Steroids were added in 5 μl ethanol, and the reaction was initiated by the addition of NADPH to a final concentration of 2-5 mM. Incubations were carried out for 10 min at 37 °C and terminated by the addition of 5 ml dichloromethane. Organic extractions and thin-layer chromatography analysis were as described above. The formation of 5α-reduced steroid products was linear with respect to protein over a 10 to 50 μg range and with respect to incubation time over a 1 to 30 minute period.

B. RESULTS

2. Identification and Analysis of Human Steroid 5α-reductase cDNAs.

To isolate clones encoding the human steroid 5α-reductase, cDNA libraries constructed from prostate mRNA were screened at reduced stringency with a radiolabeled fragment corresponding to the coding region of the rat cDNA. A total of five cDNA clones were isolated after screening 3 X 10^6 recombinants from two different cDNA libraries. Each of these cDNAs was subjected to restriction enzyme mapping and DNA sequencing and represented one species of mRNA.

The sequence of the longest cDNA insert and the predicted amino acid sequence of the human steroid 5α-reductase protein are shown in Fig. 7. The DNA sequence predicts a prostate mRNA of at least 2.1 kilobases having a 3'-untranslated region of approximately 1.3 kilobases. Within the 3'-untranslated sequence, a polyadenylation signal (AATAAA) is located 15 nucleotides 5' to a polyadenine tract, suggesting that the 3' end of this cDNA is authentic. A 5'-untranslated region of 30 nucleotides preceded a translation reading frame of 780 nucleotides encoding the steroid 5α-reductase protein.

RNA blotting experiments indicated that this cDNA hybridized to a single species of human prostate mRNA of about 2.3 kilobases. Southern blot analysis and screening of human genomic DNA libraries have similarly revealed the presence of any one functional gene homologous to this steroid 5α-reductase cDNA.

2. Structure of Human Steroid 5α-Reductase and Comparison to the Rat Enzyme

The amino acid sequence of the human steroid 5α-reductase was deduced from the cDNA insert by comparison to that of the functional rat enzyme. The human enzyme is 259 residues in length with a predicted molecular weight of 29,462. Over 40% of the amino acids in this sequence are hydrophobic, and only 16% have positively-or negatively-charged side chains (Fig. 7). These observations are consistent with the intracellular membrane location of the enzyme.

A lineup of the human and rat steroid 5α-reductase protein sequences is shown in Fig. 8. The human enzyme is four amino acids longer at the amino terminus than the rat enzyme, and surprisingly, the overall identity between these two proteins is only 60%. There is a single methionine residue in the first 89 amino acids of the human protein, whereas there are three methionines in the first 19 residues of the rat protein (Fig. 8). A comparison of the human protein sequence to those present in the Genentech Corp. and National Biomedical Research Foundation Data Bank (release 59) did not reveal any extensive homologous sequences.

3. Expression of Human and Rat Steroid 5α-Reductase in COS Cells

To determine if the observed sequence differences between the human and rat steroid 5α-reductase proteins affected their biochemical properties, the two cDNAs were expressed in simian COS cells. For the rat cDNA, a fragment corresponding to nucleotides 1 to 1975 was ligated into the pCMV4 expression vector. For the human cDNA, a fragment corresponding to nucleotides 1 to 842 of Fig. 7 was initially ligated into pCMV4. Subsequent transfection studies revealed that expression of this human cDNA yielded a ten-fold lower amount of steroid 5α-reductase enzyme activity than that obtained from the rat cDNA. Inspection of the sequence at the 5'-end of the human cDNA revealed an upstream ATG at position 5 (Fig. 7) that could conceivably result in spurious translation initiation, leading to the observed reduction in expression. To test this hypothesis, the polymerase chain reaction was used to: 1) introduce an unique Bgl II restriction enzyme site in the 5'-untranslated region of the cDNA, 2) remove the upstream ATG sequence, and 3) recreate an

optimal context (44) for the ATG of steroid 5α-reductase. Transfection of this modified human cDNA into COS cells led to the expression of levels of steroid 5α-reductase enzyme activity that equalled those obtained with the rat cDNA construct (see below).

Fig. 9 shows the results of an in vivo time course experiment in which COS cells were transiently transfected with expression vectors harboring the rat or human steroid 5α-reductase cDNAs, or with the pCMV4 vector alone. Forty-eight hours after transfection, [^{14}C]-testosterone was added to the cell media at a final concentration of 2.5 μM, and conversion of this substrate into 5α-reduced steroid products was monitored at the indicated times by thin-layer chromatography. Cells transfected with either the rat or human steroid 5α-reductase cDNAs converted half of the starting substrate into product in 1 hour (Fig. 9). The background conversion in the vector-alone transfected cells was low, with only 0.5% conversion occurring after 1 h.

This high level of expression of the cDNAs made possible the assay of steroid 5a-reductase activity in vitro in homogenates derived from the transfected cells. Homogenates were prepared as described above and various biochemical parameters were first optimized to obtain maximum steroid 5α-reductase activity. Both the rat and human enzymes demonstrated a broad pH optima centering around 7.0. The inclusion of NADPH in the COS cell homogenization buffer did not have an effect on the stability of either enzyme. The specific activities of the expressed enzymes were in the nmol/min/mg protein range and were thus equal to that reported for liver homogenates of female rats (54).

Table I shows the apparent K_m and v_{max} values determined in vitro for both the rat and human steroid 5α-reductases with five different steroid substrates. The kinetic constants were determined from a Lineweaver-Burk plot of steroid 5α-reductase activity in the presence of 0.6 to 20 μM substrate and the apparent K_m and V_{max} values were determined by linear regression analysis. The apparent K_m values obtained for these three substrates are in good agreement with those reported in the literature (55). Similarly, both enzymes demonstrated very low activities towards 11β-substituted steroids such as cortisol or corticosterone (Table I).

TABLE I

Characterization in vitro of rat and human steroid 5α-reductases expressed in transfected COS cells.

	RAT			HUMAN		
SUBSTRATE	K_m (μM)	V_{max} (nmol/min/mg protein)	K_i (nM)	K_m (μM)	V_{max} (nmol/min/mg protein)	K_i (nM)
Testosterone	2.5	1.4 - 2.5	-	3.6	0.7 - 3.6	-
Androstenedione	2.8	1.3 - 2.2	-	1.7	1.1 - 5.3	-
Progesterone	0.5	1.2 - 1.8	-	0.8	1.1 - 5.0	-
Cortisol	-	< 0.1	-	-	< 0.1	-
Corticosterone	-	< 0.1	-	-	< 0.1	-

	RAT			HUMAN		
SUBSTRATE	K_m (μM)	V_{max} (nmol/min/mg protein)	K_i (nM)	K_m (μM)	V_{max} (nmol/min/mg protein)	K_i (nM)
INHIBITOR						
4-MA	-	-	5.0,7.0	-	-	7.0,8.0
MK-906	-	-	3.0,4.0,5.0	-	-	340,380,620

COS cell transfection, cell homogenate preparation, and enzyme assay were carried out as described in Materials and Methods. Each K_m value represents the average of at least two experiments carried out on different days using cell lysates prepared from different transfections. Although both enzymes were active against cortisol and corticosterone, the amounts of 5α-reduced products formed were too small to obtain accurate kinetic constants.

The apparent K_i values was then determined for two 4-aza-substituted steroids (4-MA and MK-906) that are inhibitors of both the human and rat steroid 5α-reductase enzymes (10). Studies were initially carried out in vitro following a protocol in which two concentrations of [^{14}C]testosterone substrate were employed in the presence of increasing concentrations of a given inhibitor. The data obtained were analyzed using Dixon plots to determine the type of inhibition and the apparent K_i value (56). Fig. 10 illustrates the results

obtained in a typical experiment with extracts prepared from COS cells transfected with the human cDNA. The results for both inhibitors with the rat and human enzymes are summarized in Table I. The 4-MA compound was found to inhibit both the rat and human enzymes in a competitive fashion with an apparent K_i in the low nanomolar range, an observation in accord with previously reported values (10). Surprisingly, MK-906 was much less potent as an inhibitor of the human enzyme (K_i = 340-620 nM), than it was of the rat enzyme (K_i = 3-5 nM).

To confirm these in vitro results, IC_{50} values were determined for the MK-906 and 4-MA inhibitors using intact COS cells transfected with the human and rat steroid 5α-reductase cDNAs. As indicated in Fig. 11, both compounds were equipotent in inhibiting the rat enzyme. However, 4-MA was approximately ten-fold more potent than MK-906 in inhibiting the human enzyme.

C. Discussion

The isolation and characterization of a human steroid 5α-reductase cDNA is described by the present Example. RNA blot analysis indicated that a single species of mRNA was present in prostate. DNA sequence analysis showed that the human enzyme contained 259 amino acids that were 60% identical to those of the rat enzyme. Transfection of the human and rat cDNAs into Simian COS cells led to a high level of steroid 5α-reductase activity in both intact cells and cell homogenates. The reaction constants calculated for various steroid substrates and inhibitors revealed similarities and differences between the human and rat enzymes.

Of the differences observed between these two proteins, several are notable. First, the two enzymes are quite different in their amino acid sequences (Fig. 8). This lack of conservation is most striking in the amino terminal 130 residues in which only 50% of the amino acids are identical. A 75% conservation in the carboxyl-terminal half leads to an overall identity of 60%. With the exception of a four amino acid extension at the amino terminus of the human protein, maximum identity by alignment did not require the introduction of any gaps into the two sequences. The hydropathy plots of the human rat enzymes, as calculated by the algorithm of Kyte and doolittle (50), are almost identical. Thus, even though only 60% of their amino acids are shared, the two proteins may have retained similar secondary structures. Interestingly, at the nucleic acid level the two cDNAs are 70% identical in their coding regions, a value that is commonly derived from comparison of other rat and human cDNA homologues (57).

The biochemical behavior of the rat and human enzymes expressed in COS cells is also indicative of conservation in the presence of disparate structures. Thus, the two proteins show a preference for progesterone as a substrate over testosterone and androstenedione (Table I). Similarly, although both enzymes would reduce Δ^4, 11β-substituted steroids, measurement of activity required extensive incubation times with these substrates. The fact that similar apparent K_ms were measured for naturally occurring steroid substrates for the rat and human enzymes suggests that the two cDNAs may encode homologues and not different isozymes.

Both the rat and human enzyme activities in COS cell homogenates were most active at physiological pH values. This result was unexpected for the human enzyme as previous reports determined a pH optima of about 5.0-5.5 in cell homogenates prepared from prostate (10), epididymides (11), or genital skin fibroblasts (12). The existence of a steroid 5α-reductase activity with an alkaline pH optima has also been reported in human fibroblasts (12,58). The relationship between these acidic and alkaline pH optima enzymes and the protein encoded by the human cDNA is at present unknown. It is possible that there are two steroid 5α-reductase genes in humans; however current genetic evidence supports the existence of only one gene encoding steroid 5α-reductase in the human genome (59).

Differences in the biochemical behavior of the rat and human proteins were revealed with the use of synthetic 4-azasteroid inhibitors (Table I). Thus, the compound 4-MA was a potent competitive inhibitor of both the rat and human enzymes, while MK-906 was a 10 to 100-fold better inhibitor of the rat protein than the human protein in both in vivo and in vitro assay systems (Figs. 10 and 11). A profound difference between the inhibitory capacities of these two steroids was not detected previously in a study comparing the biochemical behavior of the rat, human, and dog prostate steroid 5α-reductase enzymes (10). The reason for this discrepancy is at present not known, but may be related to differences in the specific activities of the enzymes assayed (nmol/min/mg protein (Table I), versus pmol/min/mg protein (10), or to the different cellular environments of the two enzymes (simian kidney cells here, v rsus human prostate (10)).

The ability to express high levels of the human steroid 5α-reductase in mammalian cells should facilitate the design of more powerful and specific inhibitors of this enzyme for therapeutic use (60). In addition, the availability of the human cDNA should allow the characterization of mutations in steroid 5α-reductase that reduce enzyme activity and lead to male pseudohermaphroditism (3).

While the compositions and methods of this invention have been described in terms of preferred embodiments, it will be apparent to those of skill in the art that variations may be applied to the composition, methods and in the steps or in the sequence of steps of the method described herein without departing from the concept, spirit and scope of the invention. More specifically, it will be apparent that certain agents which are both chemically and physiologically related may be substituted for the agents described herein while the same or similar results would be achieved. All such similar substitutes and modifications apparent to those skilled in the art are deemed to be within the spirit, scope and concept of the invention as defined by the appended claims.

REFERENCES

The references listed below are incorporated herein by reference to the extent that they supplement, explain, provide a background for or teach methodology, techniques and/or compositions employed herein.

1. Wilson, J.D. (1975), Handbook of Physiology eds. Greep, R.O., Astwood, E.B. (American Physiological Society, Washington, D.C.) pp. 491-508.
2. Wilson, J.D. (1978), Ann. Rev. Physiol., 40:279-306.
3. Griffin, J.E., *et al.* (1989), The Metabolic Basis of Inherited Disease eds. Scriver, C.R., *et al.* (McGraw-Hill, New York) pp. 1919-1944.
4. Andersson, S., *et al.* (1989), J. Biol. Chem., 264:16249-16255.
5. Wilson, J.D. (980), Am. J. Med., 68:745-756.
6. Mooradian, A.D., *et al.* (1987), Endocr. Rev., 8:1-28.
7. Cunha, G.R., *et al.* (1987), Endocr. Rev., 8:338-362.
8. Brooks, J.R., *et al.* (1981), Endocrinology, 109:830-836.
9. Vermeulen, A. *et al.* (1989), Prostate, 14:45-53.
10. Liang, T., *et al.* (1985), Endocrinology, 117:571-579.
11. Fisher, L.K., *et al.* (1978), J. Clin. End. Metab., 47:653-664.
12. Moore, R.J., *et al.* (1975), J. Biol. Chem., 250:7168-7172.
13. Walsh, P.C., *et al.* (1974), N. Engl. J. Med., 291:944-949.
14. Imperato-McGinley, J., *et al.* (1974), Science, 186:1213-1217.
15. Moore, R.J., *et al.* (1973), Endocrinology, 93:581-592.
16. Jost. A. (1960), Harvey Lect., 55:201-226.
17. Page, D.C., *et al.* (1987), Cell, 51:1091-1104.
18. Evans, R.M., *et al.* (1988), Cell, 52:1-3.
19. Cate, R.L., *et al.* (1986), Cell, 45:685-698.
20. Chang, C., *et al.* (1988), Proc. Natl. Acad. Sci. U.S.A., 85:7211-7215.
21. Lubahn, D.B., *et al.* (1988), Mol. Endocrinol., 2:1265-1275.
22. Tilley, W.D., *et al.* (1989), Proc. Natl. Acad. Sci. U.S.A., 86:327-331.
23. Evans, R.M. (1988), Science, 240:889-895.
24. Griffin, J.E., *et al.* (1989), The Metabolic Basis of Inherited Disease, (Scriver *et al.* eds) pp. 1919-1944, McGraw-Hill, New York.
25. Wilbert, D.M., *et al.* (1983), J. Clin. Endocrinol. Metab., 56:113-120.
26. Moore, R.J., *et al.* (1972), J. Biol. Chem., 247:958-967.
27. McGuire, J.S., *et al.* (1960), J. Biol. Chem., 235:1634-1635.
28. Ichihara, K., *et al.* (1987), Biochem. Int., 15:1005-1011.
29. McGuire, J.S., *et al.* (1960), J. Biol. Them., 235:3112-3117.
30. Noma, Y., *et al.* (1986), Nature, 319:640-646.
31. Masu, Y., *et al.* (1987), Nature, 329:836-838.
32. Lubbert, H., *et al.* (1987), Proc. Natl. Acad. Sci. U.S.A., 84:4332-4336.
33. Julius, D., *et al.* (1988), Science, 84:4332-4336.
34. Hediger, M.A., *et al.* (1987), Nature, 330:379-381.
35. Peacock, S.L., *et al.* (1988), J. Biol. Chem., 263:7838-7845.
36. Maniatis, T., *et al.* (1982), Molecular Cloning: A laboratory Manual, pp. 1-545, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
37. Schweinfest, C.W., *et al.* (1982), Proc. Natl. Acad. Sci. U.S.A., 79:4997-5000.
38. Smith, L.M., *et al.* (1986), Nature, 321:674-679.
39. Sufhof, T.C., *et al.* (1987), Cell, 48:1061-1069.
40. Geliebter, J., *et al.* (1986), Proc. Natl. Acad. Sci. U.S.A., 83:3371-3375.
41. Walter, P., *et al.* (1981), J. Cell Biol., 91:545-550.

42. Farkash, Y., *et al.* (1988), Proc. Natl. Acad. Sci., U.S.A., 85:5824-5828.
43. Miller, W.L. (1988), Endocr. Rev., 9:295-318.
44. Kozak, M. (1986), Cell, 44:283-292.
45. Liang, T., *et al.* (1985), J. Biol. Chem., 260:4890-4895.
46. Cheng, K.C. (1988), FASEB J., 2:355 (Abstr.).
47. Cunha, G.R., *et al.* (1987), Endocr. Rev., 8:338-362.
48. Mooradian, A.D., *et al.* (1987), Endocr. Rev., 8:1-28.
49. Lippman, M.E. (1981), William's Textbook of Endocrinology, (Wilson *et al.*, eds.) pp. 1309-1326, 7th Ed., W.B. Saunders Company, Philadelphia.
50. Kyte, J., *et al.* (1982), J. Mol. Biol., 157:105-132.
51. Sambrook, J., *et al.* (1989), Molecular Cloning: A Laboratory Manual (Cold Spring Harbor Lab., Col Spring Harbor, New York).
52. Andersson, S., *et al.* (1989), J. Biol. Chem., 264:822-8229.
53. Saiki, R.K., *et al.* (1985), Science, 230:1350-1354.
54. Yates, F.E., *et al.* (1958), Endocrinology, 63:887-902.
55. Frederiksen, D.W., *et al.* (1971), J. Biol. Chem., 246:2584-2593.
56. Dixon, J., *et al.* (1979), Enzymes (Academic press, New York, NY).
57. Gonzalez, F.J. (1989), Pharm. Rev., 40:243-288.
58. Moore, R.J., *et al.* (1976), J. Biol. Chem., 251:5895-5900.
59. Imperato-McGinley, J., *et al.* (1986), Trends in Genetics, 2:130-133.
60. Metcalf, B.W., *et al.* (1989), Trends in Pharmaceutical Science, 10:491-495.
61. Adelman *et al.* (1983), DNA, 2:183.
62. Messing *et al.* (1981), Third Cleveland Symposium or Macromolecules and Recombinant DNA, Editor A. Walton, Elsevier, Amsterdam.
63. Crea *et al.* (1978), Proc. Natl. Adad. Sci. U.S.A., 75:5765.
64. Bolivar *et al.* (1977), Gene, 2:95.
65. Chang *et al.* (1978), Nature, 375:615.
66. Itakura *et al.* (1977), Science, 198:1056.
67. Goeddel *et al.* (1979), Nature, 281:544.
68. Goeddel *et al.* (1980), Nucleic Acids Res., 8:4057.
69. EPO Appl. Publ. No. 0036776.
70. Siebwenlist *et al.* (1980), Cell, 20:269.
71. Stinchcomb *et al.* (1979), Nature, 282:39.
72. Kingsman *et al.* (1979), Gene, 7:141.
73. Tschemper *et al.* (1980), Gene, 10:157.
74. Jones (1977), Geneics, 85:12.
75. Hitzeman *et al.* (1980), J. Biol. Chem., 255:2073.
76. Hess *et al.* (1968), J. Adv. Enzyme Reg., 7:149.
77. Holland *et al.* (1978), Biochemistry, 17:4900.
78. Tissue Culture, Academic Press, Kruse and Peterson, editors (1973).
79. Fiers *et al.* (1978), Nature, 273:113.
80. Thomsen *et al.* (1984), PNAS, 81:659-663. See also Boshart *et al.* (1985), Cell, 41:521-530.
81. Jobling *et al.* (1987), Nature, 325:622-625; Browning *et al.* (1988), JBC, 263:9630-9634.
82. Seeburg (1982), DNA, 1:239-249.
83. Okayama *et al.* (1983), MCB, 3:280-289.

Claims

1. A DNA segment encoding a human steroid 5α-reductase, the segment being obtainable by a process that comprises:

   (a) Preparing a clone bank made up of host cells comprising DNA segments of recombinant human DNA, at least one host cell including a DNA segment encoding a protein comprising the amino acid sequence of Figure 7;

   (b) Screening the clone bank to identify a host cell that comprises a human steroid 5α-reductase DNA segment, the clone bank being screened by a method that comprises:

   A: Preparing an oligonucleotide primer or probe comprising a stretch of the DNA sequence as set forth in Figure 7; and hybridizing the probe with DNA from said host cells to identify a host cell that comprises a human steroid 5α-reductase DNA fragment; or

B: Expressing the recombinant DNA by culturing the host cells under appropriate conditions to effect said expression; testing for the production of steroid 5α-reductase by
(1) incubating said host cells with radioactively labelled steroid;
(2) homogenizing said host cells;
(3) extracting said steroid;
(4) subjecting said steroid to thin-layer chromatography;
(5) cutting out the radioactive zones from the chromatoplates;
(6) subjecting the radioactive zones to liquid scintillation counting;
(7) identifying the products by comparison to R_F values of known standards in order to detect 5α-reduced forms of said steroid

(c) Selecting the host cell so identified and preparing the DNA segment therefrom.

2. The DNA segment of claim 1, wherein said DNA segment encodes a human steroid 5α-reductase having an amino acid sequence having more than 60% identity with the amino acid sequence of Figure 7.

3. The DNA segment of claim 2, further defined as including the steroid 5α-reductase encoding nucleic acid sequence of Figure 7.

4. A recombinant vector which includes a DNA segment in accordance with claim 1.

5. A recombinant host cell incorporating a recombinant DNA segment corresponding to the DNA segment of claim 1.

6. The recombinant host cell of claim 5, further defined as a eukaryotic host cell.

7. The recombinant host cell of claim 6, wherein the DNA segment is integrated into the genome of the host cell.

8. The recombinant host cell of claim 5, wherein the DNA segment is positioned on a recombinant vector.

9. The recombinant host cell of claim 5, further defined as capable of expressing a biologically active steroid 5α-reductase.

10. A DNA segment comprising a sequence encoding a human steroid 5α-reductase, wherein the segment comprises at least a 20 nucleotide long stretch of the DNA sequence as set forth in Figure 7.

11. The DNA segment of claim 10, wherein the segment comprises at least a 30 nucleotide long stretch of the DNA sequence as set forth in Figure 7.

12. A method for the production of steroid 5α-reductase comprising the steps of:
(a) preparing a recombinant host as defined by claim 9; and
(b) culturing the recombinant host under conditions effective to allow the production of steroid 5α-reductase by the host.

13. A method for determining the ability of a candidate substance to affect the enzymatic activity of steroid 5α-reductase, comprising the steps of:
(a) preparing steroid 5α-reductase in accordance with claim 12; and,
(b) testing the steroid 5α-reductase with a candidate substance to determine the ability of the substance to affect an enzymatic function of the steroid 5α-reductase.

14. The method of claim 13, wherein the steroid 5α-reductase is tested to determine whether it is inhibited by the candidate substance.

15. A DNA segment comprising a sequence encoding a human steroid 5α-reductas , selected from the sequence of Figure 7, a DNA sequence coding for the same amino acid sequence as the DNA sequence of Figure 7, and a sequence hybridizing to the DNA sequence of Figure 7, except a sequence which is identical to the DNA sequence of Figure 4.

16. A human steroid 5α-reductase encoded by the DNA segment of claim 15.

17. Use of the human steroid 5α-reductase of claim 16, for the identification of substances which inhibit or otherwise modify or alter the enzymatic function of the reductase.

Patentansprüche

1. DNA-Segment, welches für eine humane Steroid 5α-Reduktase kodiert, wobei das Segment durch ein Verfahren erhältlich ist, welches umfaßt:
   (a) Herstellen einer Klon-Bank, die ausgehend von Wirtszellen hergestellt ist, welche DNA-Segmente rekombinanter, humaner DNA umfassen, wobei mindestens eine Wirtszelle ein DNA-Segment einschließt, das für ein Protein kodiert, welches die Aminosäuresequenz der Figur 7 umfaßt;
   (b) Durchsuchen der Klon-Bank, um eine Wirtszelle zu identifizieren, die ein humanes Steroid 5α-Reduktase-DNA-Segment umfaßt, wobei die Klon-Bank durch ein Verfahren durchsucht wird, welches umfaßt:
      A: Herstellen eines Oligonukleotid-Primers oder einer Oligonukleotid-Sonde, welcher(s) einen Abschnitt der in Figur 7 dargestellten DNA-Sequenz umfaßt; und Hybridisieren der Sonde mit DNA von den besagten Wirtszellen, um eine Wirtszelle zu identifizieren, welche ein humanes Steroid 5α-Reduktase-DNA-Fragment umfaßt; oder
      B: Exprimieren der rekombinanten DNA durch Kultivieren der Wirtszellen unter geeigneten Bedingungen, um die besagte Expression zu bewirken; Testen auf die Herstellung von Steroid 5α-Reduktase durch
         (1) Inkubieren der besagten Wirtszellen mit radioaktiv markiertem Steroid;
         (2) Homogenisieren der besagten Wirtszellen;
         (3) Extrahieren des besagten Steroids;
         (4) Aussetzen des besagten Steroids einer Dünnschichtchromatographie;
         (5) Ausschneiden der radioaktiven Bereiche aus den Chromatographieplatten;
         (6) Aussetzen der radioaktiven Bereiche einer Flüssigszintillationsmessung;
         (7) Identifizieren der Produkte durch Vergleich mit R_F-Werten bekannter Standards, um 5α-reduzierte Formen des besagten Steroids nachzuweisen;
   (c) Auswählen der so identifizierten Wirtszelle und Herstellen des DNA-Segments daraus.

2. DNA-Segment nach Anspruch 1, worin das besagte DNA-Segment für humane Steroid 5α-Reduktase mit einer Aminosäuresequenz, welche mehr als 60 % Identität mit der Aminosäuresequenz der Figur 7 aufweist, kodiert.

3. DNA-Segment nach Anspruch 2, welches weiterhin dahingehend definiert ist, daß es die für die Steroid 5α-Reduktase kodierende Nukleinsäuresequenz der Figur 7 umfaßt.

4. Rekombinanter Vektor, welcher ein DNA-Segment gemäß Anspruch 1 umfaßt.

5. Rekombinante Wirtszelle, welche ein rekombinantes DNA-Segment, das dem DNA-Segment gemäß Anspruch 1 entspricht, beinhaltet.

6. Rekombinante Wirtszelle nach Anspruch 5, welche weiterhin dahingehend definiert ist, daß sie eine eukaryontische Wirtszelle ist.

7. Rekombinante Wirtszelle nach Anspruch 6, worin das DNA-Segment in das Genom der Wirtszelle integriert ist.

8. Rekombinante Wirtszelle nach Anspruch 5, worin das DNA-Segment auf einem rekombinanten Vektor positioniert ist.

9. Rekombinante Wirtszelle nach Anspruch 5, welche weiterhin dahingehend definiert ist, daß sie dazu fähig ist, eine biologisch aktive Steroid 5α-Reduktase zu exprimieren.

10. DNA-Segment, welches eine Sequenz umfaßt, die für eine humane Steroid 5α-Reduktase kodiert, worin das Segment mindestens einen 20 Nukleotide langen Abschnitt der in Figur 7 dargestellten DNA-

Sequenz umfaßt.

11. DNA-Segment nach Anspruch 10, worin das Segment wenigstens einen 30 Nukleotide langen Abschnitt der in Figur 7 dargestellten DNA-Sequenz umfaßt.

12. Verfahren zur Herstellung von Steroid 5α-Reduktase, umfassend die Schritte des:
(a) Herstellens eines rekombinanten Wirts, wie durch Anspruch 9 definiert; und
(b) Kultivierens des rekombinanten Wirts unter Bedingungen, die die Herstellung von Steroid 5α-Reduktase durch den Wirt erlauben.

13. Verfahren zum Nachweisen der Fähigkeit einer in Frage kommenden Substanz, die enzymatische Aktivität der Steroid 5α-Reduktase zu beeinflussen, umfassend die Schritte des:
(a) Herstellens von Steroid 5α-Reduktase gemäß Anspruch 12; und
(b) Testens der Steroid 5α-Reduktase mit einer in Frage kommenden Substanz, um die Fähigkeit der Substanz nachzuweisen, eine enzymatische Funktion der Steroid 5α-Reduktase zu beeinflussen.

14. Verfahren nach Anspruch 13, worin die Steroid 5α-Reduktase getestet wird, um zu bestimmen, ob sie durch die in Frage kommende Substanz gehemmt wird.

15. DNA-Segment, welches eine Sequenz umfaßt, die für eine humane Steroid 5α-Reduktase kodiert, ausgewählt aus der Sequenz der Figur 7 einer DNA-Sequenz, die für dieselbe Aminosäuresequenz wie die DNA-Sequenz der Figur 7 kodiert und einer Sequenz, die an die DNA-Sequenz der Figur 7 hybridisiert, ausgenommen eine Sequenz, die identisch ist mit der DNA-Sequenz der Figur 4.

16. Humane Steroid 5α-Reduktase, welche durch das DNA-Segment gemäß Anspruch 15 kodiert wird.

17. Verwendung der humanen Steroid 5α-Reduktase gemäß Anspruch 16 für die Identifizierung von Substanzen, die die enzymatische Funktion der Reduktase hemmen oder anderweitig modifizieren oder verändern.

Revendications

1. Segment d'ADN codant pour une stéroïde 5α-réductase humaine, le segment pouvant être obtenu selon un procédé comprenant:
(a) la préparation d'une banque de clones constituée de cellules-hôtes comprenant des segments d'ADN d'ADN humain recombinant, au moins une cellule-hôte contenant un segment d'ADN codant pour une protéine comprenant la séquence d'aminoacides de la figure 7;
(b) le criblage de la banque de clones pour identifier une cellule-hôte comprenant un segment d'ADN de stéroïde 5α-réductase humaine, la banque de clones étant criblée selon un procédé conprenant:
A: la préparation d'une amorce ou d'une sonde oligonucléotidique comprenant une partie de la séquence d'ADN illustrée sur la figure 7; et l'hybridation de la sonde avec de l'ADN provenant des cellules-hôtes, pour identifier une cellule-hôte comprenant un fragment d'ADN de stéroïde 5α-réductase humaine; ou
B: l'expression de l'ADN recombinant par culture des cellules-hôtes dans des conditions appropriées pour effectuer cette expression; l'essai de production de stéroïde 5α-réductase:
(1) en incubant les cellules-hôtes avec un stéroïde à marqueur radioactif;
(2) en homogénéisant les cellules-hôtes;
(3) en extrayant ledit stéroïde;
(4) en soumettant ce stéroïde à une chromatographie sur couche mince;
(5) en découpant les zones radioactives à partir des plaques chromatographiques;
(6) en soumettant les zones radioactives à un comptage de scintillations en phase liquide;
(7) en identifiant les produits par comparaison avec des valeurs R_F d'étalons connus afin de détecter les formes 5α-réduites du stéroïde;
(c) la sélection de la cellule-hôte ainsi identifiée et la préparation du segment d'ADN à partir de celle-ci.

2. Segment d'ADN selon la revendication 1, codant pour une stéroïde 5α-réductase humaine ayant une séquence d'amino-acides identique à plus de 60 % à la séquence d'amino-acides de la figure 7.

3. Segment d'ADN selon la revendication 2, défini en outre comme comprenant la séquence d'acide nucléique codant pour une stéroïde 5α-réductase de la figure 7.

4. Vecteur recombinant, comprenant un segment d'ADN selon la revendication 1.

5. Cellule-hôte recombinante comprenant un segment d'ADN recombinant correspondant au segment d'ADN de la revendication 1.

6. Cellule-hôte recombinante selon la revendication 5, définie en outre comme une cellule-hôte eucaryote.

7. Cellule-hôte recombinante selon la revendication 6, dans laquelle le segment d'ADN est intégré dans le génome de la cellule-hôte.

8. Cellule-hôte recombinante selon la revendication 5, dans laquelle le segment d'ADN est incorporé dans un vecteur de recombinaison;

9. Cellule-hôte recombinante selon la revendication 5; définie en outre comme étant capable d'exprimer une stéroïde 5α-réductase biologiquement active.

10. Segment d'ADN comprenant une séquence codant pour une stéroïde 5α-réductase humaine, le segment comprenant une partie d'une longueur d'au moins 20 nucléotides de la séquence d'ADN illustrée sur la figure 7.

11. Segment d'ADN selon la revendication 10, comprenant une partie d'au moins 30 nucléotides de longueur de la séquence d'ADN illustrée sur la figure 7.

12. Procédé de production de stéroïde 5α-réductase, comprenant les étapes consistant:
(a) à préparer un hôte recombinant tel que défini dans la revendication 9; et
(b) à cultiver l'hôte recombinant dans des conditions efficaces pour permettre la production de stéroïde 5α-réductase par l'hôte.

13. Procédé de détermination de la capacité d'une substance candidate à influencer l'activité enzymatique d'une stéroïde 5α-réductase, comprenant les étapes consistant:
(a) à préparer une stéroïde 5α-réductase selon la revendication 12; et
(b) à tester la stéroïde 5α-réductase avec une substance candidate, pour déterminer la capacité de la substance à influencer une fonction enzymatique de la stéroïde 5α-réductase.

14. Procédé selon la revendication 13, dans lequel la stéroïde 5α-réductase est testée pour déterminer si elle est inhibée par la substance candidate.

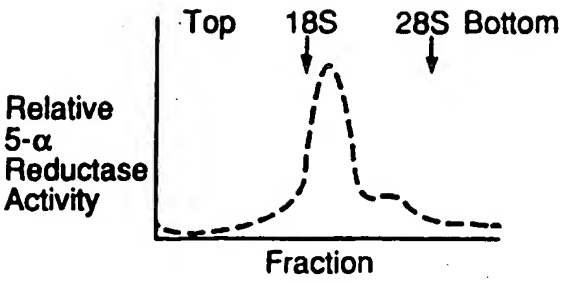
15. Segment d'ADN comprenant une séquence codant pour une stéroïde 5α-réductase humaine, choisie parmi la séquence de la figure 7, une séquence d'ADN codant pour la même séquence d'amino-acides que la séquence d'ADN de la figure 7, et une séquence s'hybridant avec la séquence d'ADN de la figure 7, à l'exception d'une séquence identique à la séquence d'ADN de la figure 4.

16. Stéroïde 5α-réductase humaine codée par le segment d'ADN de la revendication 15.

17. Utilisation de la stéroïde 5α-réductase humaine de la revendication 16, pour l'identification de substances inhibant ou modifiant autrement la fonction enzymatique de la réductase.

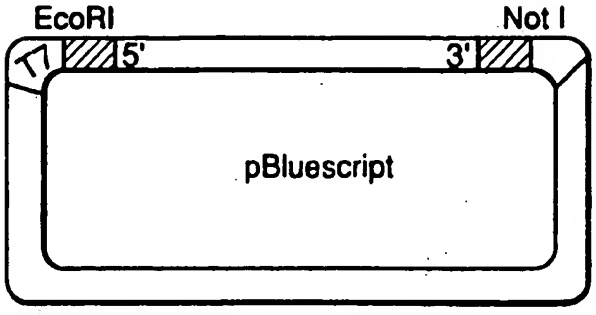
Sucrose gradient fractionation of ♀ rat liver RNA

FIG.1A



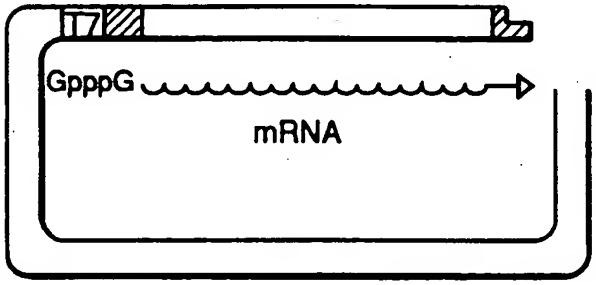
Construction of oriented, size-fractionated cDNA library

FIG.1B



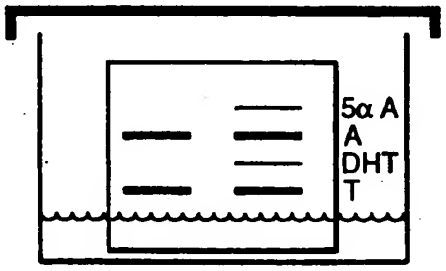
In vitro transcription of plasmid pools with T7 RNA polymerase

FIG.1C



Injection of *Xenopus* oocytes and assay of 5α-reductase activity by TLC

FIG.1D



Subdivision of positive pools to yield unique 5α-reductase cDNA

FIG.1E

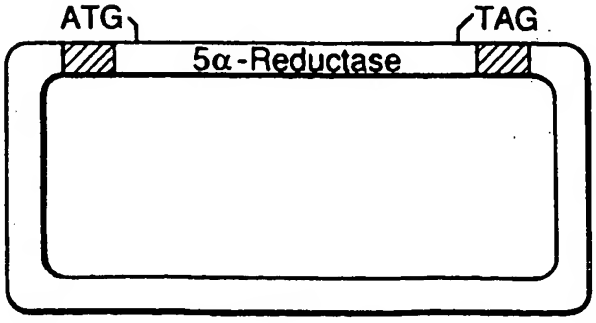


FIG.2

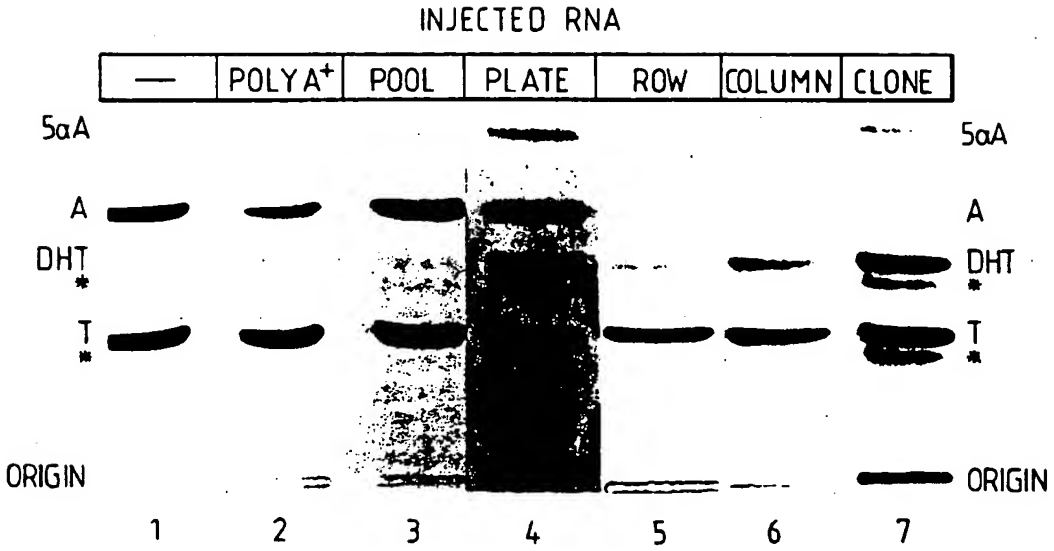

INJECTED RNA
—
POLY A+
POOL
PLATE
ROW
COLUMN
CLONE
5αA
A
DHT
T
ORIGIN
1
2
3
4
5
6
7

FIG.3

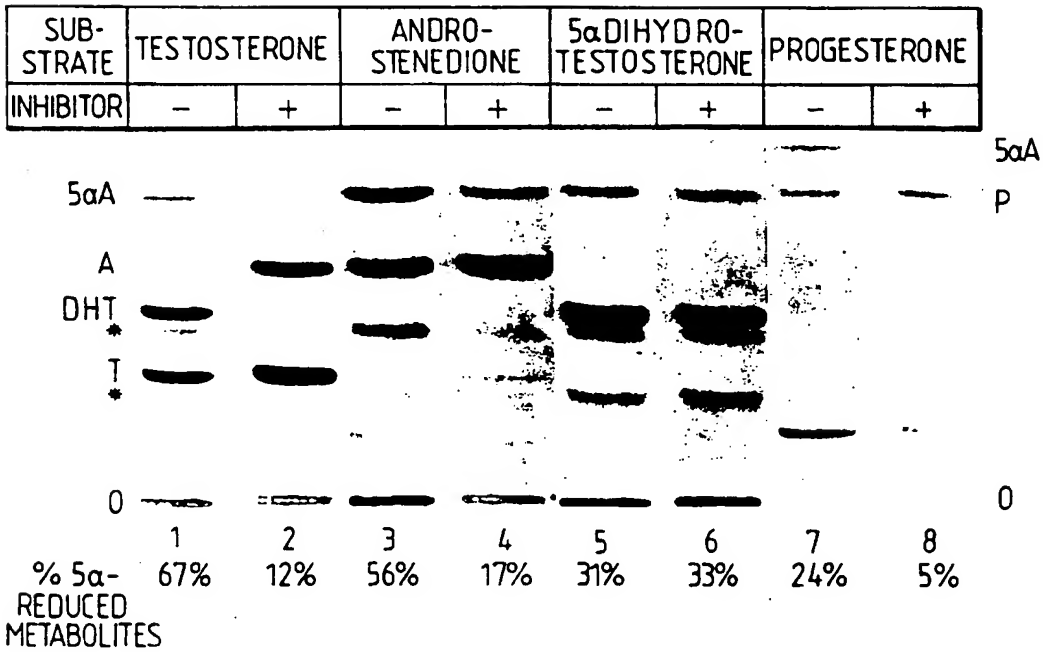

SUB-STRATE
TESTOSTERONE
ANDRO-STENEDIONE
5αDIHYDRO-TESTOSTERONE
PROGESTERONE
INHIBITOR
–
+
5αA
P
A
DHT
T
0
1
2
3
4
5
6
7
8
% 5α-REDUCED METABOLITES
67%
12%
56%
17%
31%
33%
24%
5%

FIG.4A1

EP 0 528 906 B1

FIG.4AII

```
    20                                  30
MetAlaPheValSerIleValGlyLeuArgSerValGlySerProTyrGlyArgTyr
ATGGCCTTCGTGTCCATTGTGGGGCTCCGGTCGGTTGGCTCTCCGTACGGCCGCTAC    120

    60                                  70
SerMetAlaTrpProLeuTyrGluTyrIleArgProAlaAlaAlaArgLeuGlyAsn
TCGATGGCCTGGCCGCTGTACGAGTACATTCGTCCTGCAGCCGCGCGACTGGGCAAC    240

   100                                 110
ProValLeuIleArgGlyGlyLysProThrLeuLeuValThrPheValLeuAlaPhe
CCGGTTCTGATCAGGGGAGGGAAGCCCACCCTCCTGGTCACCTTTGTCTTGGCCTTC    360

   140                                 150
GluAspTrpValThrHisProCysPheLeuThrGlyPheAlaLeuTrpLeuValGly
GAAGACTGGGTGACCCATCCCTGTTTCCTGACAGGCTTTGCCCTGTGGTTAGTGGGC    480

   180                                 190
TyrLysIleProArgGlyGlyLeuPheGluTyrValSerAlaAlaAsnTyrPheGly
TACAAGATACCCAGGGGAGGCCTGTTTGAATACGTATCTGCAGCCAACTATTTTGGG    600

   220                                 230
LeuPheThrLeuSerThrLeuLeuThrArgAlaLysGlnHisHisGlnTrpTyrHis
CTGTTCACACTCAGCACACTGCTCACCAGAGCGAAGCAGCACCATCAGTGGTACCAT    720

GTTAGCGCTGTTGCCTCCCATGAGCTGAGTCTGTCTGTCTCCCTGGTGACTTTGCTC    840

GGGGTGGGGGGGTGTCGTCCCCTGGTAAAGGACAAAGCCAATGATAAACTAATCCAC    960
```


EP 0 528 906 B1

FIG.4AIII

EP 0 528 906 B1

FIG.4AIV

```
CGCTCTCTGACTTCAGCAATCAGCAGCCCTTACAATCCTGCAAGATTCCACCCAAGT   1080
GCCTGGAGGTGTGGGAGTGTGGCCTAAGCACAGTCTGCCATCCTTGACCGCAGACCT   1200
CCTCAGGAGGTGAGACAGCAGAGTGCTTCCATTCACTCGATGACCCCATTTTTGCTC   1320
AAAACAAAATATCAAAACAAAATTTGTTATTTTGAATGCACCCAAGGACCAATCATG   1440
AATGAAGATAACAACCAGCTAACTGTCCAAAGAAGTGACCGCAGTAATAAAAGACGC   1560
CTGGGCAGGCTGAGGCCCACTGGTGAAGAGCCATTCCCACCGGACCCATGCACACTG   1680
TAGAACTGTGTGCTCTGAACCCAGTGAAGATGCTGCATGAAGACCTGCAGGCACACC   1800
CTCTGTGTGCCCTTTCCAGCTGGCTTTCCCATCAGGGCTTCCTCAGCTCTTCTGCTC   1920
TCATGACACAGCTACTTCATTTTAACAAAGAGCAGTGTTTAATGGGGAACTACCCTT   2040
GTAGGTCTCCTCTCAAAACCTCAGGCCTGTCTGGTGTGTTCTGAAACGTTTGTGTGG   2160
TGGTGCTAGATAAAGTTGGAACCTAGGACTCCAGGTTGCTAGGCGGATGCCCTGACA   2280
ATACTAAGTGAGATTAATTTAAGAGGAATCCTGTCCTAACACTGTATACTTCATTCC   2400
GTAAAAA...  2470
```


EP 0 528 906 B1

FIG.4B

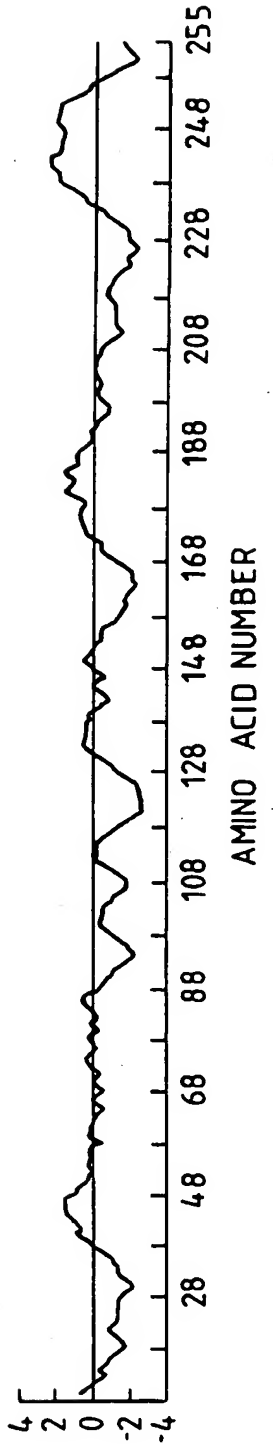
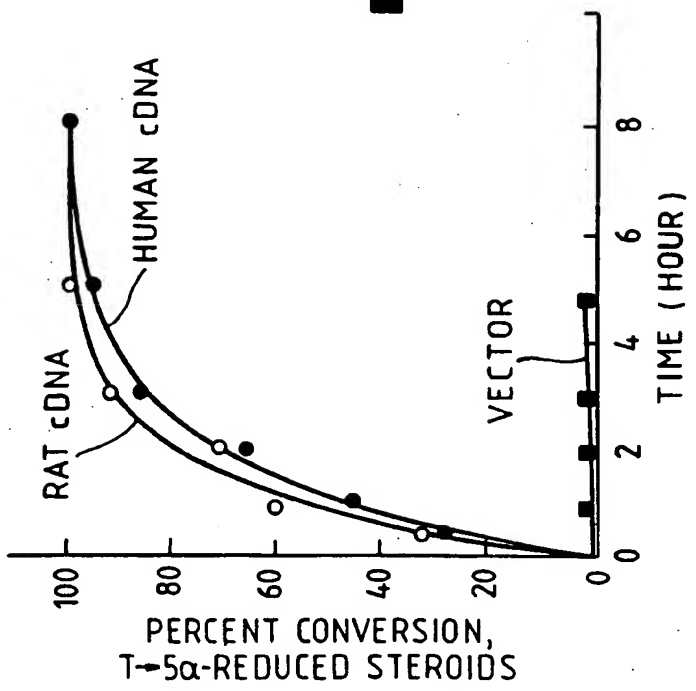

4
2
0
-2
-4
28
48
68
88
108
128
148
168
188
208
228
248
255
AMINO ACID NUMBER

FIG.9


RAT cDNA
HUMAN cDNA
VECTOR
100
80
60
40
20
0
0
2
4
6
8
PERCENT CONVERSION,
T→5α-REDUCED STEROIDS
TIME (HOUR)

5α-R RNA	−	−	+	+	+	+
MICROSOMES	−	+	−	+	+	+
TRYPSIN	−	−	−	−	+	+
TRITON	−	−	−	−	−	+

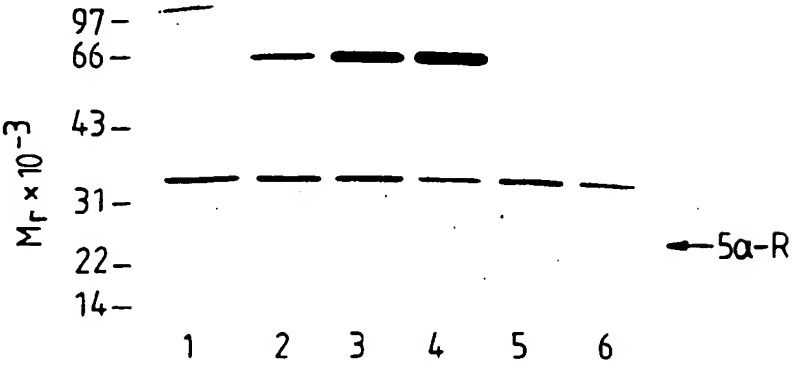


FIG.5

FIG.6A

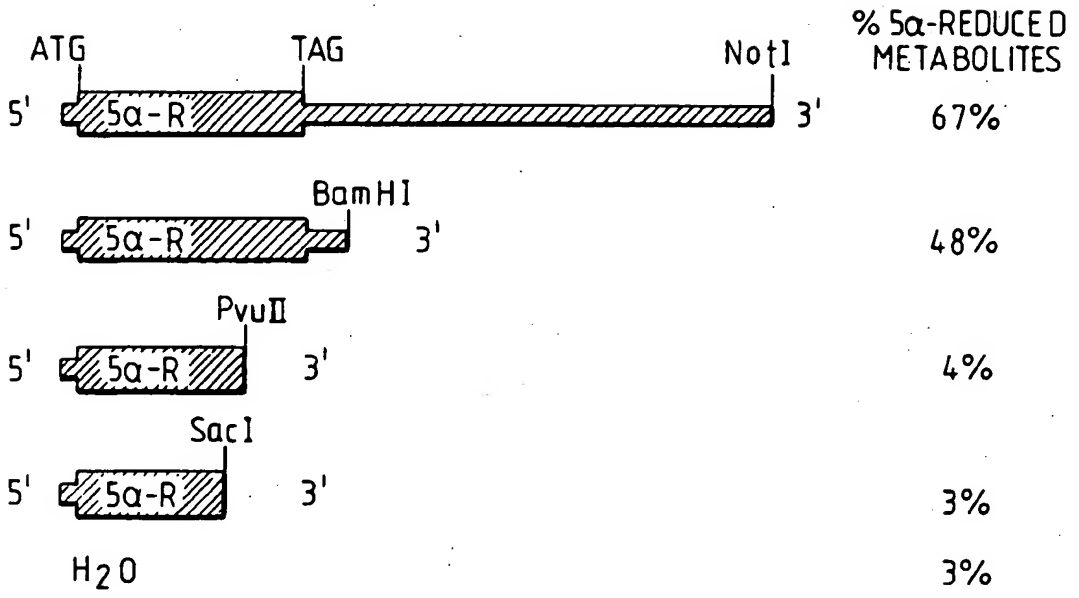

% 5α-REDUCED METABOLITES
ATG
TAG
NotI
5'
5α-R
3'
67%
BamHI
48%
PvuII
4%
SacI
3%
H2O
3%

FIG.6B

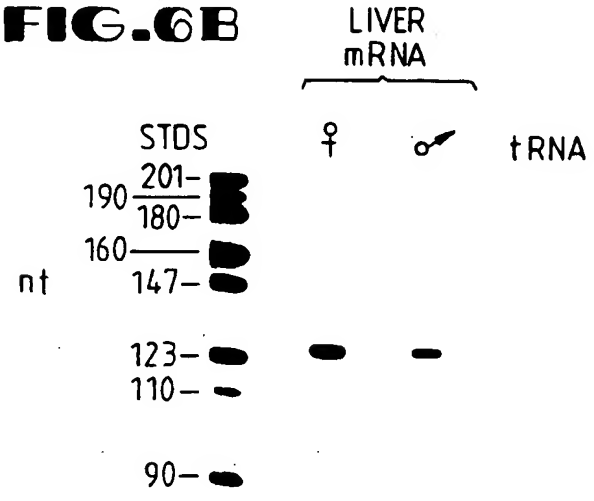
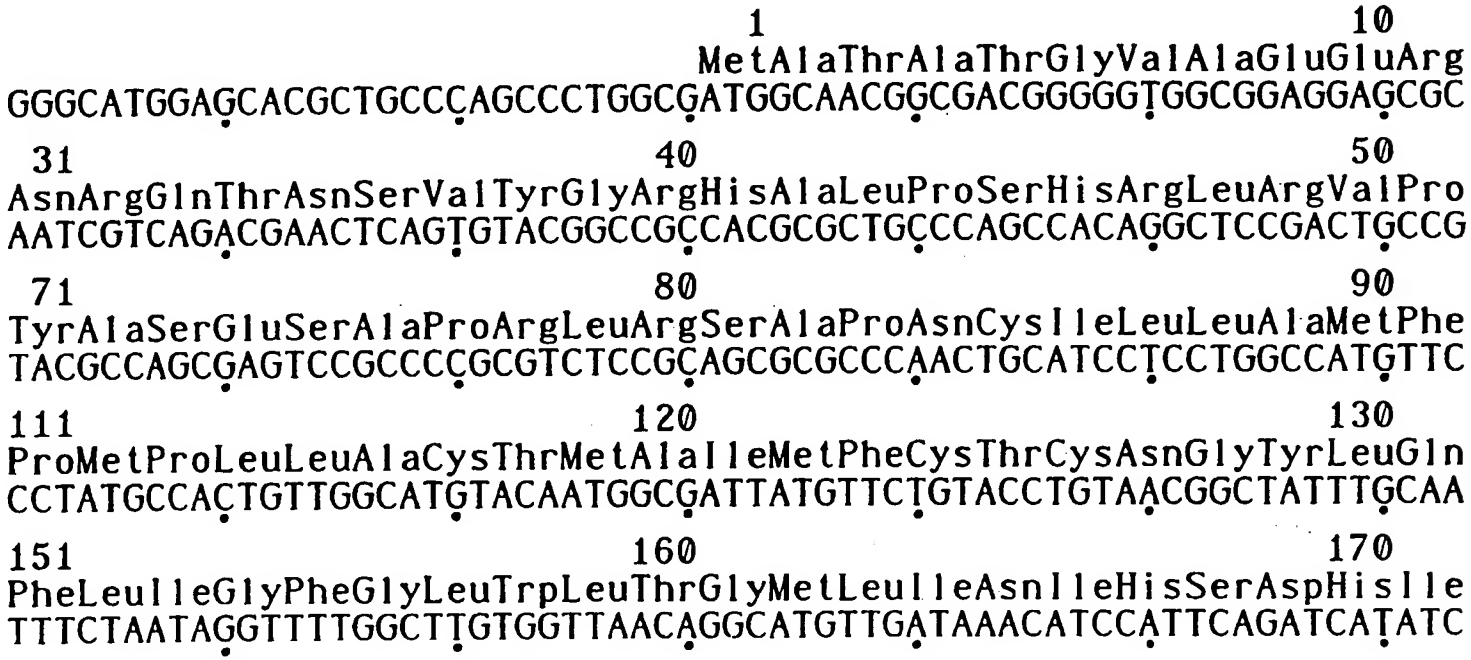
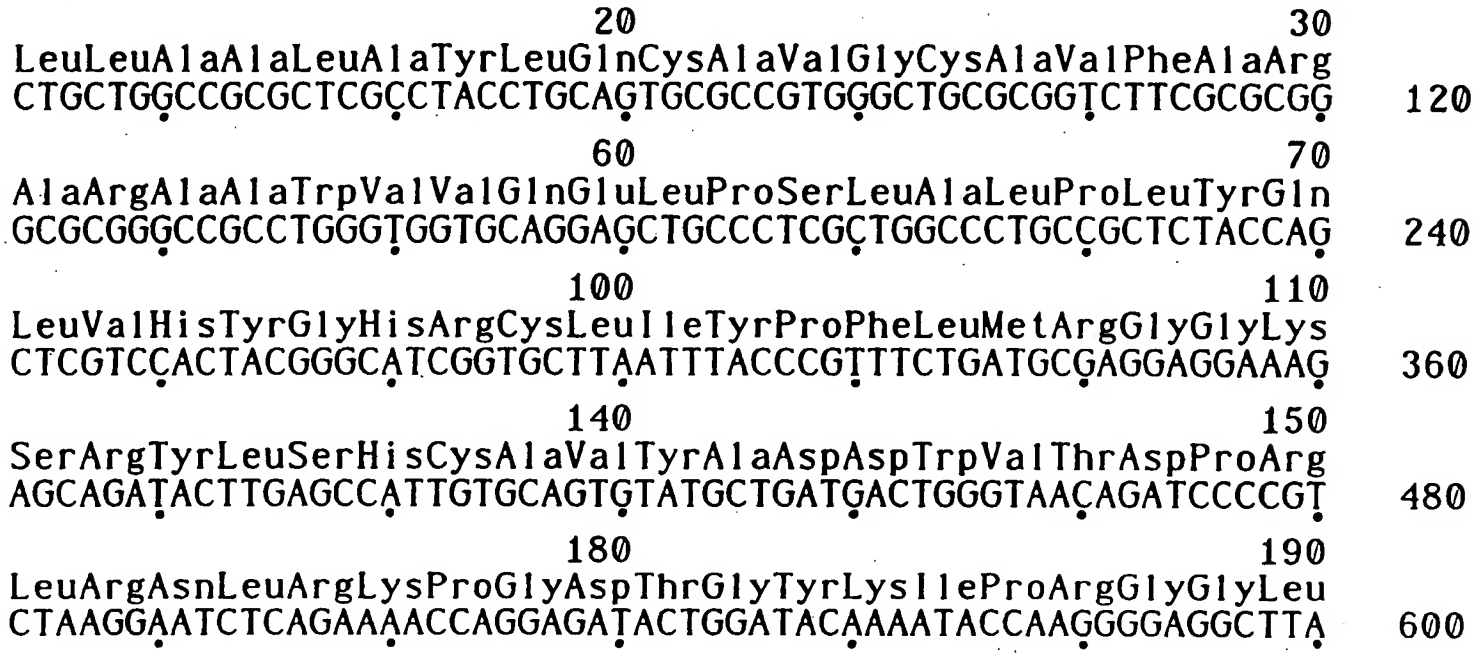

LIVER mRNA
STDS
♀
♂
tRNA
nt
201–
190
180–
160
147–
123–
110–
90–

FIG.7A



EP 0 528 906 B1

FIG.7B



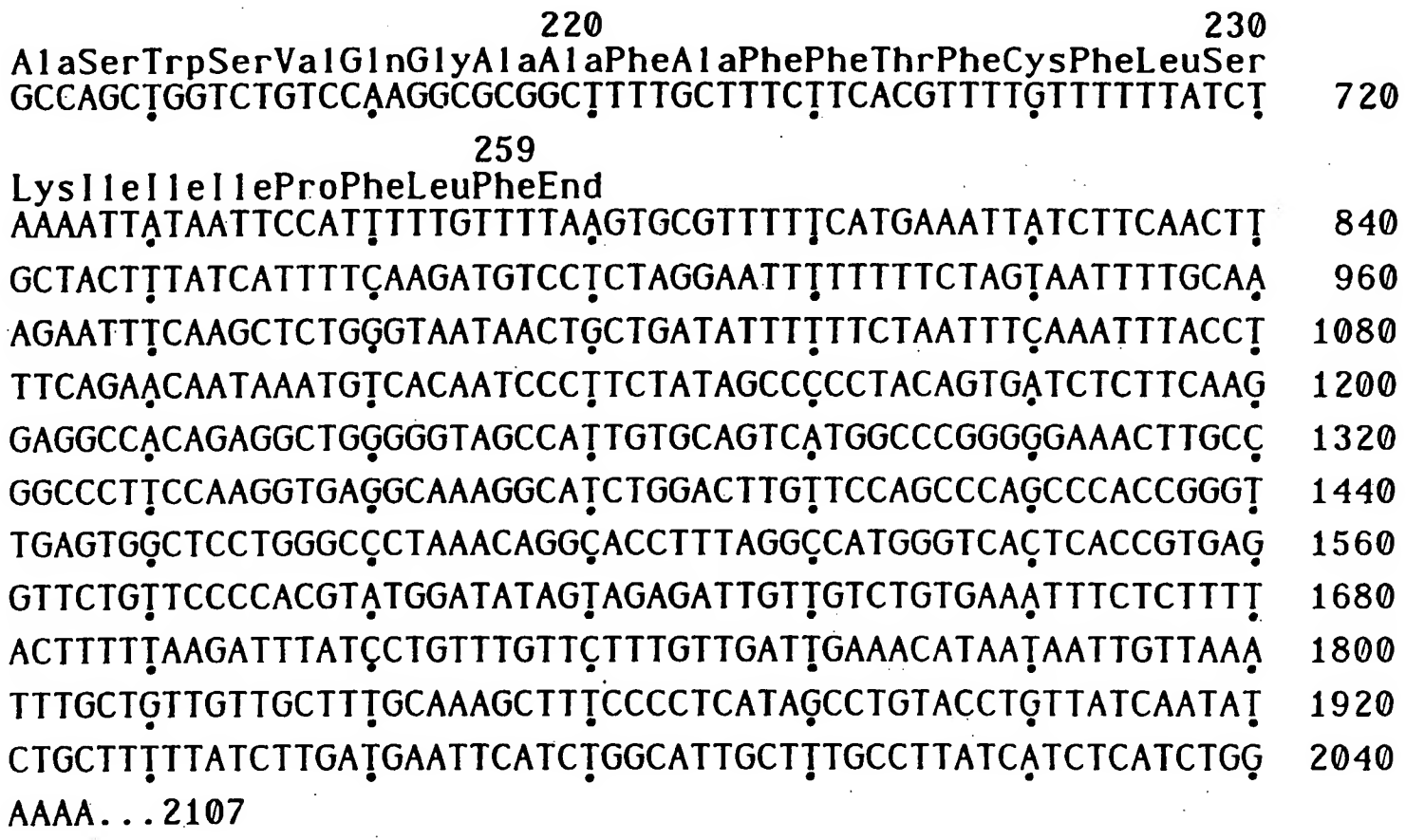
EP 0 528 906 B1

FIG.7C

```
191                                200                                210
PheGluTyrValThrAlaAlaAsnTyrPheGlyGluIleMetGluTrpCysGlyTyrAlaLeu
TTTGAATACGTAACTGCAGCCAACTATTTTGGAGAAATCATGGAGTGGTGTGGCTATGCCCTG

231                                240                                250
GlyArgAlaLysGluHisHisGluTrpTyrLeuArgLysPheGluGluTyrProLysPheArg
GGTAGAGCAAAAGAGCATCATGAGTGGTACCTCCGGAAATTTGAAGAGTATCCAAAGTTCAGA

GAAGCTTTCCAATGGCGCTTCTCTATGGACTTTGTAAATAAGTTATATCTTTGTAATTTTCCT

TCTACCTAATAAGTACCTAAATACGCTGAAATGGAGGTTGAATATCCTACTGTGTAACAGGTC

CTTTTGGCTATGTCTTGCCAAGTGTGTATGAGACTAGACTTTACAACTGTCTTTGATGGCATT

GTCAACTGCAGTGTTGCTTCCCTCCCCCTATAGGGCTGGAATCTGTCTAGGAGCCCTCTCTCG

AACCTTCGTGTCAGGTGCTGTGTGTAAGTGGAGAACTTGGGGATAGAGGAGGAAGCTCCTCGT

GACATCACCGGGCAGGGAGGGGTGCTGGTGGTGGTTCATACGGAGTAAGCTGCTCTGCCTGTG

CCATCAATGTGCTCTGGTCTGACATGGTTTCTCTCTGTCTTCTAGTCTAGACCTAGTTTTTTT

GTAGATTTTGAGTTTTCCCTTGTAGTGTAAAGAATGATCACTTTCTGTAACAATAACAAGACC

ATTCTCTACAGCCTTCTTTTTCTTCCATAGCTAATCTTCCTTCTAATAGTTTTTGCTTTCTGT

AAAATAATCTTCCTGTTGAATGCTTCATGACTTGAATTCTACTTTGATAAAAACATTGCCATA

AGTTTTTAAATGCCATTTGTTTCAGTTGTCTTTAACAACATAATAAATAGACTTTGCCATTTA
```


EP 0 528 906 B1

FIG.7D



EP 0 528 906 B1

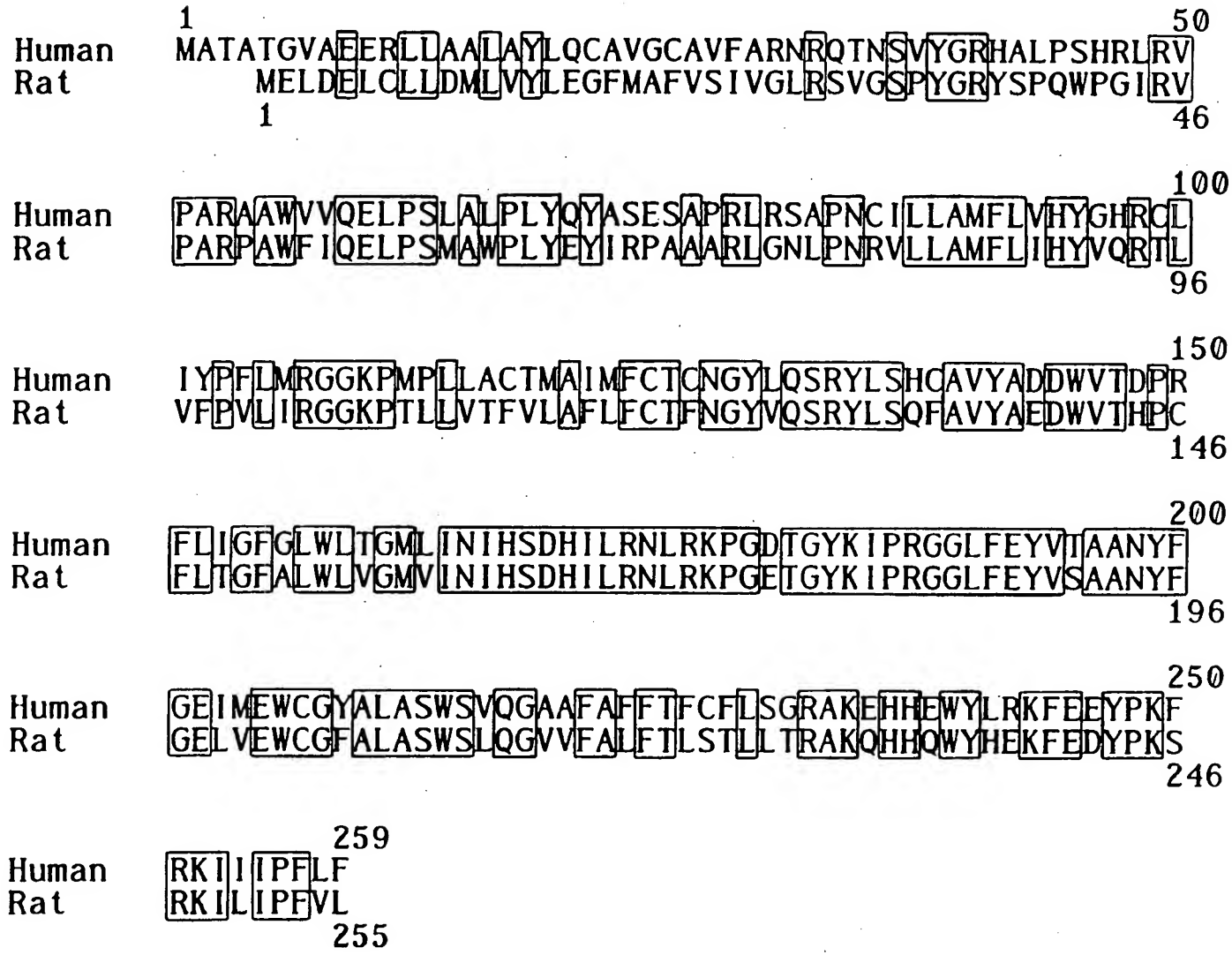


FIG.8

EP 0 528 906 B1

FIG.10A

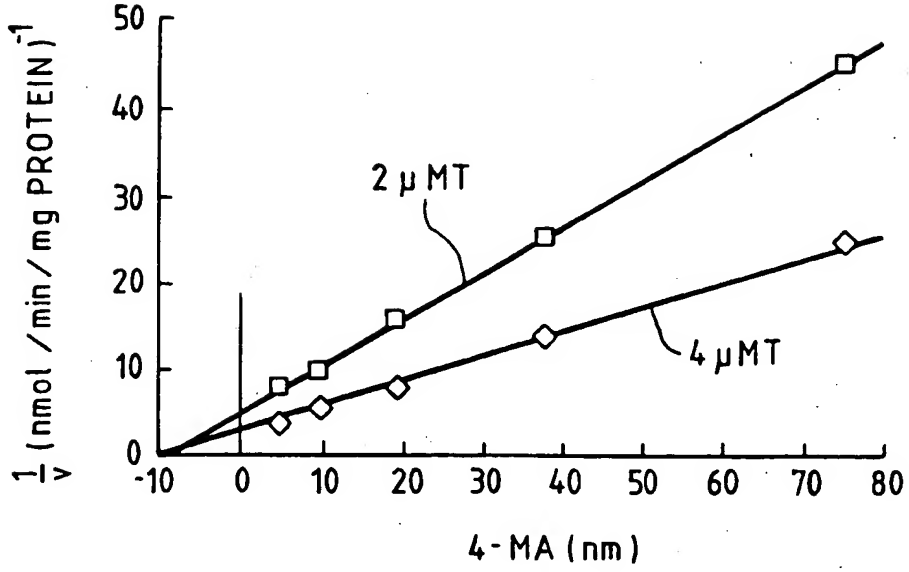
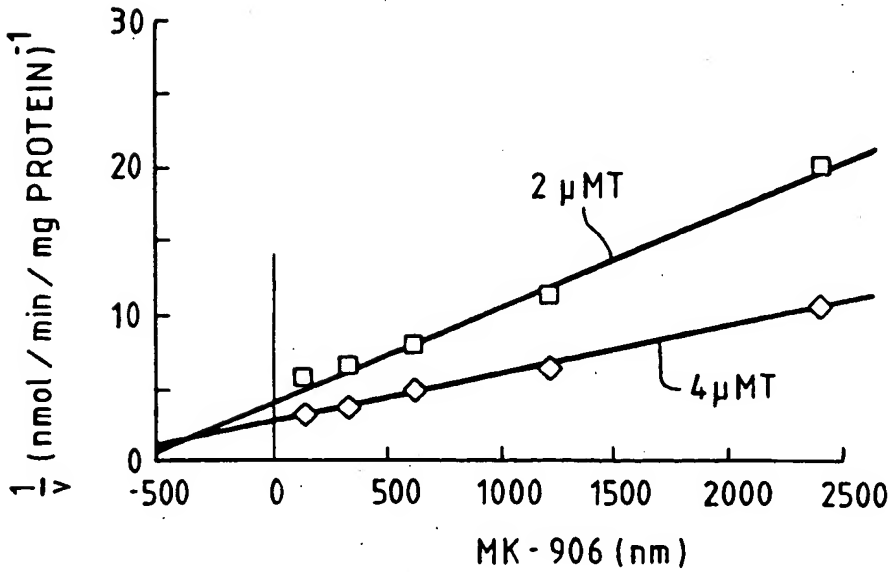


FIG.10B



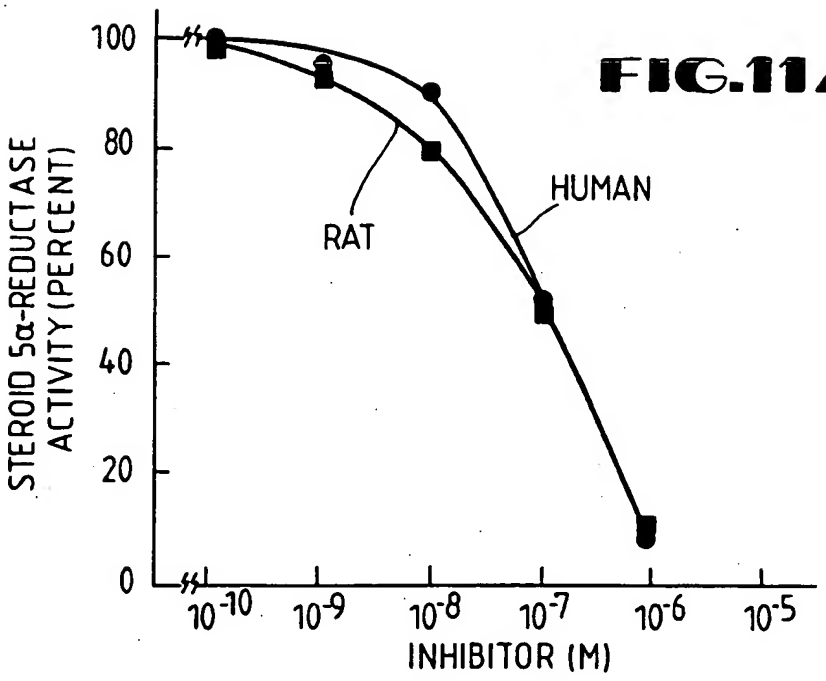


FIG.11A

FIG.11B